OE FORM 6000, 2/69

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
OFFICE OF EDUCATION

ERIC REPORT RESUME

ERIC ACC. NO.: ED 042 069

CH ACC. NO.: AA 000 607

P.A.:

PUBL. DATE:

ISSUE: RIEJAN71

IS DOCUMENT COPYRIGHTED? YES ☐ NO ☒

ERIC REPRODUCTION RELEASE? YES ☐ NO ☒

LEVEL OF AVAILABILITY I ☒ II ☐ III ☐

AUTHOR: Castner, Myra H.; And Others

TITLE: Granite School District First Grade Reading Study.

SOURCE CODE: YEX29710

INSTITUTION (SOURCE): Granite School District, Salt Lake City, Utah

SP. AG. CODE:

SPONSORING AGENCY:

EDRS PRICE: 0.75;6.50

CONTRACT NO.:

GRANT NO.:

REPORT NO.:

BUREAU NO.:

AVAILABILITY:

JOURNAL CITATION:

DESCRIPTIVE NOTE: 128p.

DESCRIPTORS: *Reading Instruction; *Grade 1; *Teaching Methods; *Hypothesis Testing; *Comparative Analysis

IDENTIFIERS:

ABSTRACT A comparative study of first-grade reading instructional methods was undertaken with the support of the Granite School District Exemplary Center for Reading Instruction. This study was conducted in 19 schools of the district and involved approximately 1,295 students. Nine hypotheses concerning the various approaches used in reading instruction were tested. The report of this study is divided into five chapters: (1) Introduction, (2) The Experimental Program, (3) Methods and Procedures, (4) Statistical Results, and (5) Summary and Conclusions. Seven appendices give additional information. (CK)

GPO 870-390

N-X

ED 042 069

Granite School District
FIRST GRADE READING STUDY

AA 000 607

U.S. DEPARTMENT OF HEALTH, EDUCATION
& WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Granite School District

FIRST GRADE READING STUDY

Myra H. Castner
Dr. Gabriel Della-Piana
Michael Hogben
John E. Allen

TABLE OF CONTENTS

PAGE

ABSTRACT v

CHAPTER

I. INTRODUCTION 1
Statement of the Problem 1
The Hypotheses 2

II. THE EXPERIMENTAL PROGRAMS 5
"Linguistic" Reading Programs 5
Harper & Row *Linguistic Readers* 6
S.R.A. *Basic Reading Series* 7
The Programmed/Linguistic Approach 8
Initial Teaching Alphabet (i/t/a) 9
Multi-Sensory Approaches 10

III. METHODS AND PROCEDURES 13
Population 13
General Experimental Design 14
Teacher-Training and Program Fidelity Control 17
General Administrative and Testing Procedures 19
Instruments and Schedule of Testing 20
Statistical Treatment of Data 26

IV. STATISTICAL RESULTS 28
Relationships Between Pre-test Measures and Reading Achievement 28
Relationships of Pre-test and Achievement Measures to Pupil/Class Characteristics and Secondary Achievement Scores 35
Evaluation of Comparative Program Effectiveness by Readiness Level 38

V. SUMMARY AND CONCLUSIONS 62
  Summary of Design and Procedures 62
  Statistical Summary and Conclusions 64
  Limitations of the Study 67
  Educational Implications 71

REFERENCES 73
APPENDIX A 77
APPENDIX B 92
APPENDIX C 95
APPENDIX D 105
APPENDIX E 110
APPENDIX F 113
APPENDIX G 117

ABSTRACT

Within the last few years, an upsurge of interest in, and debate about, beginning reading methods has culminated in a wide variety of new reading instructional materials and approaches. Although each new program has been field tested, as well as supported by the research of independent investigators, differences in research design have limited possible comparisons and generalization of results. The Cooperative First Grade Study (Bond and Dykstra, 1967), although maintaining control of the instruments used, found interproject comparisons limited in validity due to the presence of factors, individual to each school district, which would interfere with the inter-district generalization of results. Therefore, it remains the responsibility of school administrators to provide for comparative research of reading programs within their own districts. Procedures and evaluation instruments may then be controlled across a variety of selected programs and comparison made between groups representative of the district population.

The reported design and conduct of a comparative study of first-grade reading instructional methods was supported by the Granite School District Exemplary Center for Reading Instruction (ECRI) with the cooperation of the Bureau of Educational Research, University of Utah. Undertaken in nineteen schools of the district and involving approximately 1295 students within 47 classrooms, it extended throughout the 1966-67 school year. A complete report of the design including statistical data and tables is available through ECRI. Primary concern was with the design of single-project research which would enable a school district to assess multiple reading programs proposed for use within that district. Due to the preliminary nature of the study and the imbalance in the number of subjects within programs, the statistical results, though important, are of secondary concern.

Following the research recommendations of Campbell and Stanley (1963), a quasi-experimental non-equivalent control group design was used in which existing classroom units, kept intact as experimental or control groups, were administered pre-tests and post-tests. Specified factors, apt to determine internal or external validity of the research, dictated procedures within this framework. Validity factors of particular importance to this study concern effects resulting from intrasession history, instrumentation, statistical regression, and biases resulting from selection.

Eighteen of the participating first-grade classrooms were control groups, randomly selected for use in establishing a norm for district first grades. Basal reading series, typical of those in general district use, served as instructional materials within these classrooms. The twenty-nine experimental classrooms were essentially those for which a linguistic, programmed-linguistic, or multi-sensory modified linguistic reading program had been provided prior to the design of this study. In an attempt to assess the multi-sensory contribution, classrooms combining the multi-sensory materials with basal reading programs were also included. The experimental groups used the Webster Division, McGraw-Hill *Programmed Reading* materials, SRA *Basic Reading Series,* Harper & Row *Linguistic Readers, Listen, Look, Learn* (LLL) materials, and Educational Developmental Laboratories (EDL) machines and materials to supplement basal reading series.

As a factor in the internal validity of the research design, fidelity to experimental programs was assigned primary importance and controlled through a workshop-classroom observation-workshop cycle, repeated during the school year. The experimental teachers were trained in pre-instructional workshops which were held with the aid of program authors, editors, or publisher representatives. On two consecutive occasions each month, each classroom, experimental and control, was observed by a research staff member trained in that program and was rated for fidelity on scales developed by the research staff. Following these observations, a workshop was held to discuss program fidelity and any problems which had arisen. In order to control the Hawthorne effect, essentially equal treatment was accorded all groups throughout the research period. Upon recommendation of linguistic publishers, a bibliography of supplementary reading materials which maintained regular sound-spelling relationships was also constructed for use in avoiding contamination of reading experiences with irregular words. A time log sheet was developed for use in computing differences in mean reading-instructional time between classrooms and a five-level socio-economic scale was also developed for use as a dependent variable in the analysis of the data.

The *Murphy-Durrell Reading Readiness Analysis* and the *Lorge-Thorndike Intelligence Test* were selected for administration as pre-tests during the second and third week of the school year. Skill-oriented pretests were administered in February during a mid-year series of tests. The *Gates-MacGinitie Primary A,* Form I, and the *San Diego Inventory of Reading Attitudes* were administered to the entire population. A written composition was also uniformly assigned to the entire population to ob-

tain measures of number of words attempted, sentence complexity (derived by use of the Subordinate Clause Index), and the mean "minimal terminal unit" (T-unit) length. A sub-sample of each classroom was given the *Gilmore Oral Reading Test* to assess fluency and word attack differentials, and observation of the selection of reading as a free-time activity was undertaken in randomly selected classrooms. Post-tests, administered in mid-May, included an alternate form of the Gates-MacGinitie group achievement test and the individual oral reading test, a second written composition, and administration to a random selection of control and "linguistic" classrooms of an experimental *Linguistic Reading Test,* presently available in only one form.

All tests were administered by the research staff following training by the project director. Enough test personnel were used to complete each series in a brief time that would avoid differences in the amount of teaching received by the groups prior to testing, and time of day was equalized among the programs. All scoring was done by trained members of a scoring staff.

With the fall from status of a one-for-all approach to the selection of reading programs, it was felt necessary to seek information concerning the relative effectiveness of reading programs with students of varying characteristics. Therefore, a twofold statistical emphasis was selected for this study. First, the Pearson product-moment correlation coefficient was utilized to determine the relationships between pupil/class characteristics, readiness, and achievement measures within each separate program, and six 21 X 21 matrixes resulted. Additionally, an investigation was made of possible differences in effectiveness among programs, with pupils grouped homogeneously by two of these factors, assessed readiness and IQ. A 3 (levels) X 6 (treatments) analysis of variance was performed separately for readiness and IQ levels for each of the twenty-one dependent variables of the study.

Succinctly, examination of pre-test/post-test correlations found the Murphy-Durrell total test to be the variable with the highest relationship to post-instructional achievement. It was, however, only slightly higher than one of its component parts, the Letter Names sub-test. This finding has pertinent implications for both readiness training and the possibility of prediction of first grade reading achievement through the measurement of knowledge of letter symbols and names alone. The Lorge-Thorndike DIQ relationships to achievement were lower than either of these readiness measures, although the DIQ was the most highly related contributing factor in readiness performance. Low to moderate relationships

to both pre-instructional readiness and subsequent achievement were found for the measures of written composition. In general, the highest correlations with achievement were found for the number of words written, while the complexity and length of constructions were relatively unaffected by the growth of first grade reading skills. The extremely high relationships found across all programs between the "linguistic" and traditional achievement measures indicate that it is unnecessary to administer a special test for "linguistically" trained groups. While a variation in familiar reading vocabulary does exist between the programs, adjustment to the unfamiliar seems to be made with equal facility by both groups. Only low correlations with both readiness and achievement measures were found for all other variables.

Blocking by readiness or IQ, no single reading program was found to be significantly better than all others in all respects or to be uniquely effective for students of any given level of pre-instructional readiness or IQ. Significant differences involving a program tended to occur across all levels and without clear discrimination among variables. McGraw-Hill *Programmed Reading* was favored most frequently. Results favoring SRA and Harper-Row were considerably less frequent but did occur, particularly in the middle and low levels. Although the extremely low *n* of the SRA and Harper-Row groups lessens the confidence in their results, they were not, with some accountable exception, surpassed significantly by the McGraw-Hill program and must, in this study, be considered essentially equal to it in performance.

The low programs in the resulting significant differences were quite consistently the Control, EDL-Basal, and Listen, Look, Learn groups. The performance of the traditional basal programs is indicative of a need for in-depth revision of these programs. Merely adding to existing programs was not evidenced to be sufficient or even profitable. On all levels, the Control basal groups were only better than similar basal programs in which time was taken for the mechanized drill of EDL materials relatively uncoordinated with the programs. The generally low performance of the multi-sensory, mechanized LLL program, which also only surpassed EDL-Basal, gives rise to question of the merits of mechanizing instruction at all. However, there would be need to investigate the potential performance of the "modified linguistic" aspect of this program, without the multi-sensory devices, before definitive conclusions could be reached.

The non-equivalent control group design, using intact classrooms, was felt to have been effectual and a useful framework for the control of factors threatening research validity. The procedures of this study for

assuring research validity were evaluated and areas of concern determined to be principally attributable to the necessity to utilize an existing selection of experimental programs and teachers. The random assignment of teachers to both experimental and control programs, rather than the use of "self-selected" experimental volunteers, and a balanced number of classrooms per program would have increased both internal and external validity. However, a recognized need to narrow the scope of future research through carefully stated, and highly specific, hypotheses is felt to be of prime importance. The time for comparative studies which broadly hypothesize "achievement" or "fluency" would seem to be past. Equally important is the selection of experimental programs which, through a contrastive analysis of similarities and differences, have been determined to allow for the testing of the stated hypotheses. There is otherwise, as evidenced in this study, a lessening of the sensitivity of obtained measures, and aspects of opposing programs are prone to overlap and cloud the results. While the procedures explored in this study may be useful in future research, the instruments selected, or developed, for use in the study will not necessarily be of value. Measurement of dependent variables is closely related to the stated hypotheses of any one research project, and instruments are necessarily selected, or developed, on this basis.

An alternative to the straight comparative study is discussed in a separate publication of the Exemplary Center for Reading Instruction, Granite School District, titled "Research Strategies for Maximizing the Effectiveness of Programmed Reading" (Della-Piana and Hogben, April, 1968). That report describes how a comparative study may be used as a preliminary step in selecting a treatment (curriculum package) which warrants evaluation and developmental work to increase its effectiveness for a given population. It discusses some research strategies for gathering data to facilitate installation, monitoring and continuous revision of a curriculum package.

CHAPTER I

INTRODUCTION

The reported design and conduct of a comparative study of first-grade reading instructional methods was supported by the Granite School District Exemplary Center for Reading Instruction (ECRI). Undertaken in nineteen schools of the district and involving approximately 1295 students within 47 classrooms, it extended through the 1966-67 school year. Primary concern was with the design of single-project research which would enable a school district to assess multiple reading programs proposed for use within that district. Due to the preliminary nature of the study and the imbalance in the number of subjects within programs, the statistical results, though important, are of secondary concern.

Statement of the Problem

Within the last few years, an upsurge of interest in, and debate about, beginning reading methods has culminated in a wide variety of new reading instructional materials and approaches. Although each new program has been field tested, as well as supported by the research of independent investigators, differences in research design have limited possible comparisons and generalization of results. As a result, a major cooperative first grade research effort was sponsored by the United States Office of Education (USOE). Through a cooperative effort it was hoped that the problems and incomparability of piece-meal research could be avoided. Twenty-seven separate projects, held together by a coordinating center located at the University of Minnesota, were conducted across the country during the 1964-65 school year. A wide variety of newly proposed reading programs and methods were evaluated, generally in comparison with basal reading programs. The Cooperative First Grade Study (Bond and Dykstra, 1967), although maintaining control of the instruments used, found inter-project comparisons limited in validity due to the presence of factors, individual to each school district, which would interfere with the inter-district generalization of results. Therefore, it remains the responsibility of school administrators to provide for comparative research of reading programs within their own districts. Procedures and evaluation instruments may then be controlled across a variety of selected programs and comparison made between groups representative of the district population.

In May of 1966, when preliminary reports of the cooperative study appeared in *The Reading Teacher* (Stauffer, 1966), it was determined that it would be appropriate for Granite School District to prepare for a longitudinal study of reading programs currently in use, or proposed for use, in its classrooms. This pilot study was initiated as a first grade research project of the district's newly established, USOE funded, Exemplary Center for Reading Instruction. It was intended principally to design comparative research for "linguistic" reading programs. However, two new multi-sensory, mechanized approaches, already planned for introduction in the district, were included in the study because of their use of "modified linguistic" materials and because of the information about design which inclusion of diversified programs might provide. A total of six new reading approaches, in use in twenty-nine district classrooms, were available as experimental groups for the study. Eighteen first grade classrooms, using materials traditional to the district, were randomly selected to establish a district reading performance norm for this grade level.

The Hypotheses

For the evaluation, a study was designed which was directed toward determination of the relative effectiveness of the new approaches as compared to each other and to the district norm performance. After consideration of area influenced by reading instruction, instruments were selected, or constructed, for assessment of their value in testing the following null hypotheses:

1. There will be no significant difference among approaches to reading achievement as measured by a standardized reading achievement test.
2. There will be no significant difference among approaches to oral reading fluency or word attack methods as measured by individual oral reading tests administered to a random sampling of each classroom.
3. There will be no significant difference among approaches to control of written expression as determined by a syntactic evaluation of unaided written composition.
4. There will be no significant difference among approaches to attitude toward reading as assessed by an attitude survey.

5. There will be no significant difference among approaches to interest in reading as a free-time activity as assessed by an activity-selection observation.

6. There will be no significant difference among approaches to effectiveness for children of different ability levels.

7. There will be no significant difference among approaches to effectiveness for children of different readiness levels.

8. There will be no significant difference among approaches to effectiveness for the instruction of boys or girls.

9. Teacher directed instructional time will have no significant relationship to any of the preceding hypotheses.

The hypotheses to be tested were made as null statements for several reasons. Foremost is the inability to predict, by any but purely intuitive means, the comparative levels of performance of the groups in the measurement of any given dependent variable. While research utilizing standardized achievement instruments can be found for most of the new approaches, differences in study design are present to prevent their use as predictive for our groups. Frequently, no differences or conflicting differences have been reported. Additionally, some of the instruments were developed for this study and would be difficult to relate to results obtained from use of even similar devices.

Effort was made, in constructing the hypotheses, to consider a wide range of educational goals for reading instruction. It was necessary to be arbitrary in the selection of the goals included, since the time available for this study imposed limitations, and agreement in skill priority, or even descriptive terminology, does not yet exist. Elements of the most commonly agreed upon "achievement" skills, which may be broadly defined as word recognition, word attack, comprehension, and reading rate, are measurable through standardized group and individual tests. However, since it can be argued that reading instruction should influence more than just the reading "act," additional hypotheses were needed. Since, "achievement" remaining constant, any one reading method might produce a better attitude toward, or interest in, reading, these hypotheses were considered significant outcomes which should be tested. If we see for children a need to better manipulate, and communicate with, a variety of language forms, then control of written expression is important for evaluation across reading methods. The variable of teaching time devoted to aspects of the reading program was included as being essential to evaluation of each tested hypothesis.

As a subsidiary purpose of this study, consideration of the proposition that different approaches to initial reading instruction might produce differentiated results in terms of the intelligence or readiness of the child was deemed to be important. While intelligence and reading readiness may not be clearly definable factors in reading success or failure within approaches, they do have the advantage of being readily available, through test instruments, for future predictive use. As almost a single voice, educators reject the premise that there can be a one-method-for-all approach to reading instruction. Yet, it is the rare classroom, school, or perhaps even district, in which more than one basic reading series is in use. It is possible that a basis for within classroom differentiation of approach is feasible through predictive use of one or more of the factors to be tested in this study.

CHAPTER II

THE EXPERIMENTAL PROGRAMS

Several aspects of the experimental design hinge upon elements found in the philosophy or techniques of the reading programs involved. It is, in fact, the variance in these elements among programs that justifies the designing of comparative research. It is, therefore, considered necessary to provide an orientation to those programs concerned in the pilot study as a foundation for discussion of the research design.

New approaches to reading instruction are most frequently viewed in relation to "traditional" basal reading programs. Although they have received increasing criticism in the last few years, basal readers do serve as the predominant materials for elementary reading instruction. Judgment has centered largely on their exclusive use of word-frequency lists for vocabulary selection, sentences controlled by length rather than structure, and limited story content, particularly in the pre-primers and primers. Satires involving the typical "Oh, Oh, Oh! Look, Dick, Look!" of pre-primers are popular with stand-up comics and educational columnists. Charge has also frequently been made of too heavy reliance on the building of sight vocabulary without adequate provision for phonic word attack skills. Sheldon (1965) has attempted to answer these criticisms with a comprehensive statement of basal programs and their related research. He has also predicted changes for the near future which include increased early emphasis on phonics, more mature story content, and the near abolishment of vocabulary control. Represented within our control groups are five of the six basal series currently approved for use in Granite School District (American Book, Allyn & Bacon, Ginn, Houghton-Mifflin, and Macmillan). Examination of the teacher's manuals of these series indicates that they include most of the limited range of variations in sequencing and emphasis available within traditional basal programs and thus serve as excellent comparative programs. A special phonics program, *Speech to Print* (Durrell and Murphy, 1964) was in supplementary use in approximately one-third of our control classrooms.

"Linguistic" Reading Programs

While the majority of the published materials in use in the experimental classrooms of this study are said to be "linguistically" conceived, it must be pointed out that this is in no way a "pure" term. As with many educational labels, even "phonics" or "whole-word" approaches, a wide range of practices is subsumed under the general rubric of "linguistic reading program." If linguists, who are often mistakenly given credit as a group for the conception of these materials, were consulted, it is doubtful that more than a small handful could arrive at agreement as to any one basis for such a reading program.

It was Leonard Bloomfield (1942) who first allied linguistics and reading methodology over twenty years ago in his essay entitled "Linguistics and Reading." He approached reading instruction through the system of English phonology or sound. The written symbol as a representation of speech sounds and the basically, though not perfectly, alphabetic nature of our writing system are the two basic principles which serve as foundation for his system of reading methodology. He introduced the concept that, when viewed in units larger than the individual letter used in "phonic" approaches, the English language contained largely "regularly" spelled words. He distinguished them from "irregularly" spelled words by their grapheme-phoneme (symbol-sound) correspondence. He based his beginning reading approach on disconnected words and nonsense syllables which presented these regular spellings in an orderly manner. This was felt to permit the inductive realization of the grapheme-phoneme patterns of our language, testable through the use of nonsense syllables. Irregularly spelled words were to be presented only after overlearning of the first group. Thus, stop, look, come, do, and go, with their multiple *o* sound correspondence, would not all appear as initial vocabulary to confuse the child and convince him of the absence of regular relationships. Content and meaning were considered unessential to beginning reading since a child was felt to be too busy with the mechanics of reading at this stage to absorb content.

Harper & Row Linguistic Readers

A review by a distinguished linguist, H. L. Smith (1963), reiterated what he considered some of the most serious defects of the Bloomfield approach. Although he accepts the principles of systematic presentation of grapheme-phoneme relationships, he believes the original materials have not been re-evaluated in terms of current knowledge of the phonemes and dialects of American English. He feels the ordering in terms of difficulty was based on the "educated guess" of a linguist rather than on research. He objects to the "divorcing of meaning" from the reading process and to the lack of illustrations throughout the materials.

Smith concluded his review with linguistic textbook recommendations of his own which have since been implemented in the Harper and Row *Linguistic Readers* (Stratemeyer and Smith, 1965), in use in one experimental classroom of this study. Familiar words which maintain a one-to-one grapheme-phoneme correspondence are presented as first words, repeated in spaced repetition in a meaningful context which employs the informal style of children. Once an initial vocabulary of highly regular words is developed in the pre-primers, additional vocabulary is developed rapidly by annexation, omission, or substitution of letters to words previously learned. Since today's children have come to expect well illustrated books, he has stressed the inclusion of pictures which illustrate rather than tell the story. Until recently, this series has only been available for first grade use. However, materials to be used through the third grade are now published and would lend themselves to a longitudinal comparative study.

SRA Basic Reading Series

Possibly the most influential restatement of the Bloomfield phonologically-based principles has come from another eminent linguist, Charles C. Fries (1963; 1965; 1966). He has expanded some concepts and added new terminology, but the basic conception of reading methodology remains much the same. Basic to his "linguistically sound approach" to reading is the fundamental relationship of oral language to reading—the only difference being in the medium through which stimuli contact the nervous system. Language is conceived of as a "code of signals" which are the means by which meanings are communicated, rather than the meanings themselves. The process of learning to read is then "the process of transfer" from auditory signals to new visual signs. After initial teaching of the alphabet symbols by name, to avoid their hindering the process of learning to read, emphasis is on the "spelling patterns" of our language, which are to be presented in orderly listed sets. They represent, but extend, Bloomfield's "regularly" spelled words and are presented in much the same contrastive letter sequences which are organized to identify the highly variable vowel phonemes of English. In agreement with Bloomfield, there are to be no pictures in the initial reading materials to distract from the task at hand and act as oftentimes misleading crutches. Once the "transfer stage" has been mastered to the point of automatic response habits, the reader moves to the "productive reading stage" during which he responds to meanings without the apparent use of signals. The last stage of "vivid imagination realization" develops an ability to read literature interpretively and constitutes a never ending process of learning to read.

The *SRA Basic Reading Series* (Rasmussen and Goldberg, 1965), in use in two experimental classrooms of this study, follows quite closely, and originates from an initial attempt to utilize, the Fries recommendations. For teaching the "transfer stage," linguistic readers are intended for use with beginning readers through mid-second grade. An initial alphabet book, which teaches the symbols by name, is followed by six readers. The one-to-one spelling-sound correspondence and "patterned" presentation of vocabulary is maintained throughout with only 72 "exceptional" words being introduced for reasons of story content. Unlike available Fries materials (Fries, Wilson, and Rudolph, 1966), rapid pacing is emphasized to avoid the possibility of only sight-learning of words and to promote the "inductive realization" of patterns. Also deviating from Fries' recommendations, pictures are used throughout the series. In the first books, they only decorate while in the others they relate to, but do not tell, the story. Unlike many "linguistic" reading programs, which limit themselves to only the "transfer stage" at present, an additional segment of this series, the *SRA Comprehensive Reading Series* for grades 2-6, is available and lends itself to a longitudinal comparative study by providing materials for the "productive reading stage." Workbooks are provided which stress the development of an ability to read increasingly complex sentence patterns.

The Programmed/Linguistic Approach

Linguist M. W. Sullivan applied his knowledge of the system of English phonology to the development of another of the experimental programs of this study, the Webster, McGraw-Hill *Programmed Reading Series* (Sullivan and Buchanan, 1963). The indispensable contributions of linguistic science to the programmed presentation of initial reading instruction are quite evident when the necessity of presenting sequenced "bits" of information is considered. A rather thorough analysis of the language as it is spoken and graphically represented would be required to produce such sequencing. In this reading series, initial vocabulary is presented in minimal pairs which vary only in one phoneme (e.g., pin-pan, an-ant, man-fan) rather than being confined entirely to spelling pattern lists. Each letter is assigned only one sound and this sound-symbol relationship is maintained well into the series to a point at which the pupil is thought to have perceived the code and be able to generalize it. Because of the tendency of vowels in unaccented syllables to be "schwaed" (e.g., the *a* in about or the *e* in remain which have a kind of "uh" sound), the first few books are limited to single syllable words.

The rationale for programming instruction began with Sidney Pressey's invention, over forty years ago, of a multiple choice teaching machine. Pressey's technology has been described by Skinner (1961) as primarily a testing device to be used after teaching had taken place elsewhere. In this early form, machine teaching had little impact upon education. However, recent general interest in programmed instruction in its current forms has resulted from the efforts of Skinner and his applications of the experimental analysis of behavior. He developed a teaching stratagem which is best described by the stimulus-response learning model. Programming involves a process of determining the objectives of a particular instructional program (Mager, 1962) and then sequencing steps, in bits of information, that will lead to the terminal objectives. Each step (frame) is a stimulus-response unit, presented to the student one at a time and requiring an active response which is promptly reinforced. Since programmed materials are individualized and self-pacing, the frustration of waiting for, or holding back, classmates is largely avoided.

The *Programmed Reading Series* is a "linear" program in which the student follows a linear, step-by-step, sequence through a consumable reading workbook. He moves from frame to frame, actively making written response and receiving reinforcement through direct and immediate knowledge of the "correctness" of his response. The program is said to have been designed to assure success at least 95 percent of the time. Each new concept is presented in a variety of ways, paired with already conditioned stimulus. It is repeated several times in different contexts to prevent boredom and ensure mastery, retention, and stimulus generalization. Three series of supplementary hard-back story books, coordinated with the instructional workbooks, are available to aid in the transfer of skills to content reading materials.

Initial Teaching Alphabet (i/t/a)[1]

While not often classified among "linguistic" reading programs, the Initial Teaching Alphabet (i/t/a) is another attempt to regularize English grapheme-phoneme relationships in beginning reading instruction. Sir James Pitman's Augmented Roman Alphabet is utilized to provide a separate graphic symbol for each phoneme of the language. Of the forty-four characters used, twenty-four are retained from the traditional alphabet.

---

[1] The single i/t/a classroom of this study served as a pilot project for a multiple-class i/t/a study which was conducted in Granite School District during the 1967-68 school year. The reader is referred to the results of the latter study, available as a separate publication of ECRI.

The twenty augmentations were designed, with transfer in mind, to preserve high visual compatability with traditional orthography (TO). It is basically a phonemic rather than phonetic alphabet. Only the most significant phonemic distinctions are transcribed exactly and traditional orthographic spellings are retained wherever possible (e.g., both c and k are retained and the unstressed schwa is represented by its traditional vowel symbol). The i/t/a is not an attempt toward either spelling or alphabet reform in the traditional sense but is intended only to be a consistent transitional medium for simplifying initial reading instruction.

There is no specific method of instruction implied in i/t/a. Methodology is a product of the materials selected for use. In England, i/t/a was originally superimposed upon the traditional basal reading programs through change in the print of available series. The more recent English *Downing Readers* (Downing, 1963), which were created specifically in i/t/a, retain the tight sight vocabulary and repetition control of eclectic methods and provide an inductive approach to the realization of sound-spelling correspondences. The American *Early-to-Read i/t/a Program* (Tanyzer and Mazurkiewicz, 1966), in use in the one experimental i/t/a classroom in this study, is a combination language experience/phonic approach. Traditional vocabulary control is virtually eliminated in the series materials. Individual sounds and symbols are taught and blended for both reading and early independent creative writing. Many major publishers duplicate their children's literature in this orthography for use as supplementary reading materials in this program. Change to TO is to be made after competence and confidence in reading is felt to be assured in i/t/a (usually toward the end of first grade).

The major comparative investigations of i/t/a instruction have been directed by authors of i/t/a series. Downing (1964) reported the results of a British longitudinal study which indicated significantly superior performance among children using i/t/a materials. At the end of a one-year study in Bethlehem, Pennsylvania, Mazurkiewicz (1966) reported results favorable to the i/t/a experimental group. However, the results of this study after two years (Stewart, 1967) were generally inconclusive. Tanyzer and Alpert (1965; 1966), Mazurkiewicz (1966), and Hahn (1966) all reported favorable i/t/a results within their individual projects participating in the cooperative first grade study. The need for longitudinal follow-up studies is recognized by these researchers and is being undertaken at present by Tanyzer and Alpert.

Multi-Sensory Approaches

The advent of "linguistic" reading programs, however loosely definable, has predictably led to the more recent publication of "modified linguistic" reading programs. The principles of one such program, the "structural" reading approach (Stern and Gould, 1965), have been combined with a variety of mechanized reading devices in the multi-sensory *Listen, Look, Learn* (LLL) program of the Educational Developmental Laboratories (EDL) in use in four of the experimental classrooms of this study. It is predominantly a mechanized program, using textbook materials (Stern and Gould, 1963) correlated in instructional principles and sequencing, only to supplement the materials for machine instruction.

In this approach, the child is to learn to read through insight into the structure of words. It begins with the spoken word and systematically progresses to written structures. Words are taught by component parts larger than the individual letter, and reading and writing are taught together. A basic difference that separates this approach from the Bloomfield-based linguistic programs is the grouping of words by beginnings (e.g., ca-n, ca-t, ca-p) rather than endings (e.g., m-an, f-an, c-an). It is felt that this technique follows natural pronunciation patterns (we say "can" as ca-n), directs the eye in desired left-to-right direction, and is more apt to promote attack of less regular, multisyllable words (e.g., calendar or camel will be tried more confidently because the beginning is known). After words are "deciphered" structurally, instant recognition is promoted through repeated practice with the word.

The program utilizes a wide variety of machines to present materials based on structural reading concepts. The Tach-X is a tachistoscope which is designed to develop high speed visual accuracy, efficiency, and retention. The Aud-X is an audio-visual machine which utilizes and synchronizes both sight and hearing in building oral language, sight vocabulary, and structural analysis skills. The Controlled Reader projects visual story materials at controllable rates through a moving slot. It directs the students' attention and promotes concentration by making rereading impossible. Use of this range of "multi-sensory" instruments is felt to provide training in skills which are pertinent to the reading process.

In order to attempt assessment of the contribution of the mechanized devices in the *Listen, Look, Learn* reading program, eight classrooms, in which the EDL machines and their materials were used to supplement basal reading programs, were included as experimental groups. While Spache (1963) has reported increases in reading rate as a result of use of mechanized devices, Tinker (1967), in a summary of research in this area, has cast doubt on the value of their use. The Tach-X and Controlled Reader and their materials were used during brief daily periods to augment the basal program. No attempt was made to correlate the somewhat diverse instructional philosophies of the basal and structural approaches. The Aud-X was not available for use in these classrooms.

CHAPTER III

METHODS AND PROCEDURES

Influences impinging upon the design and procedures of this study have been many and varied. Foremost has been the willingness of the Granite School District personnel to support the study so fully and to cooperate without complaint in all that was asked of and from them. An additional positive influence has come from the availability, for selective use and examination, of the techniques and instruments of the many recently reported comparative studies. A negative limitation has been imposed by lack of time for desirable preplanning and orderly selection of experimental programs and classrooms. However, since this has been a pilot study producing only preliminary data, these limitations will be acknowledged as unavoidable.

Population

At the time of pre-testing, 1295 first-grade students in forty-seven classrooms of Granite School District participated in the study. Mortality due to moves from these classrooms was relatively small and was balanced across programs. No teacher, either control or experimental, declined to take part and the original selection of classrooms remained in the study throughout the school year. However, the data from one classroom, excluded from all lists of experimental groups presented in this report, were withdrawn as a result of inability to maintain program fidelity to more than a small degree. Only children from this original pre-test population were included in a three-level analysis of variance utilizing pre-test readiness and IQ. When blocking by these levels, a severe imbalance occurred from the single i/t/a classroom, and this classroom's data too, was withdrawn from the present study.

The experimental and control groups ranged from the extreme district high socio-economically to the extreme district socio-economic low represented in Title I schools. A five-level socio-economic scale was developed with the aid of the district Pupil-Personnel office and was utilized as a dependent variable in analysis of the data. Factors influencing a classroom's position on the scale included average area income, average area home valuation, percentage of area students completing high school, percentage of area students predicted to attend college, and their school's status in the Title I and Headstart federal programs.

In order to establish a norm for district first grades, eighteen control classrooms were randomly selected, using a statistical table of random numbers, from schools not involved in experimental first-grade programs. Statistical procedures for random selection were adhered to rigidly with the following exceptions. The entire first-grade population of two schools in lower socio-economic areas was included as a balance for experimental programs being undertaken in area Title I schools. One other first-grade group was added at the request of the principal whose other first grades were all involved in experimental programs. The possibility of contamination of the data due to these exceptions to random selection is recognized to be important and it will be recommended that any future studies allow no such exception.

The most successful procedure for selection of experimental classrooms, following the experimental design recommendations of Campbell and Stanley (1963), would entail an equally random selection and assignment from district first-grade teachers. For this study, however, the experimental programs include those requested by either the principals and/or teachers involved and already planned for use prior to the design of the study. The only exception, the SRA materials, were assigned to the entire first-grade population of a school which had no classes selected during the random sampling for controls. The resulting experimental classrooms include twelve using Webster Division, McGraw-Hill *Programmed Reading* materials, two using the SRA *Basic Reading Series*, one using Harper & Row *Linguistic Readers,* one using i/t/a materials, five using *Listen, Look, Learn* materials, and eight using EDL machines and materials to supplement a basal reading series. The resulting imbalance, particularly in programs represented in only one classroom, further limits the interpretation of data resulting from the study but does offer maximum preliminary data.

General Experimental Design

Campbell and Stanley (1963) have outlined alternative designs for research in teaching in terms of control of specified factors which might interfere with research validity. Design of this study is based upon their "quasi-experimental non-equivalent control group design." Their use of the descriptive term "quasi-experimental" is intended only to designate a degree of variance from optimal experimental conditions and not to rule out the value of this design. Within the design, use is permitted of existing classroom units rather than the preferred random assignment of children to control or experimental groups. The "self-selection" of the experimental programs by teachers is also acceptable within this design, though recognized as imposing limitations and certainly not desirable. Broadly, the design involves experimental and control groups which do not have "pre-experimental sampling equivalence" but are given pre-tests and post-tests. Within this framework, specified factors threatening both internal and external validity must be controlled.

To prevent their direct effect on experimentation results, apt to be mistaken for experiment-related changes, seven factors of *internal* validity were specified by Campbell and Stanley and accounted for in the internal design of this study. The factor of "intrasession history," events which may occur between measurements in addition to the experimental variables, was assigned prime importance and dealt with in several ways. In an attempt to partially nullify the Hawthorne effect (Cook, 1962), which is found to bias toward the experimental groups, all participating groups were addressed as "experimental" and received essentially equal attention and treatment. Any exceptions occurred in teacher training and will be discussed in that section which follows. Fidelity to the assigned program was controlled through pretraining of teachers and classroom observation, with resulting additional teacher training, by the research staff. Experimental occasions, such as visitations or testing, were equal and randomized in respect to time of day or day of week. A computer-scorable log sheet was constructed on which teachers recorded the amount of time during a day they devoted to the direct teaching of specified reading activities. It was used in statistically accounting for differences in instructional time.

The "instrumentation" factor—changes in the measuring instruments, the observers, or scorers which may produce changes in the measures obtained—was controlled by the use of fixed (printed) tests, trained staff testers, and rigidly followed test procedures. Trained scorers, separate from the testing staff, followed detailed written scoring instructions. Since observations were used only as a training tool rather than a measurement device, they did not affect the instrumentation factor.

"Statistical regression" and biases resulting from "selection" are the factors felt by Campbell and Stanley to be the greatest threat to this particular research design. However, it was pointed out that both are solvable through recommended statistical procedures which include the use of class means and an analysis of covariance. However, the limited number of classrooms in most programs prohibited the use of class means in this study. Also, pre-test scores were used for separate three-level analyses of variance, because of the information concerning the effectiveness of each program for children in each level such a technique might provide.

Control of the remaining three factors is considered to be inherent to the basic design itself. The factors of "maturation," a function of "passage of time per se," and "testing," the possible effect of one test on a later test, are considered to be manifested equally on all groups. The "mortality" factor, possible sample bias resulting from unequal loss of subjects from the groups to be compared, was controlled by use of intact classroom units. Loss due to family moves would tend to balance out and any differential would occur only by chance.

The most frequently expressed doubts about comparative educational research are judgments about the *external* validity or generalizability of the research findings. Campbell and Stanley isolated factors which tend to jeopardize external validity in this particular design. The "interaction effects of selection bias and the experimental variables" have been partially controlled by the use of a number of classroom units. This allows for a wide variety of the conditions and schools represented in the district population to which results might be generalized. The "self-selection" of the experimental groups weakens this control to some extent but does not invalidate it completely.

The "reactive effects of testing and the experimental situation," the Hawthorne effect which we have endeavored to spread equally among all groups, will interfere with generalization to groups in a non-experimental situation. However, since a balance of these reactive effects between experimental and control groups has been attempted, the trends of the results of the study may then be generalized, with the caution due preliminary data, to Granite District first-grade classrooms.

While questions of validity controlled the overall research design, the selection and sequencing of instruments used within this framework was necessarily controlled by the hypotheses to be tested. The purpose of the pilot study, stated more fully in the hypotheses presented earlier, was to assess selected language arts skills and attitudes following initial reading instruction with diverse materials and methods. Since pre-instructional reading achievement is largely nonexistent among first graders, readiness for instruction and measurable intellectual potential were determined to be significant areas for the pretesting of the population. A group reading-readiness test and a group non-verbal IQ test were administered during the second and third week of the school year. Reading and writing skills were considered sufficient for administration of skill-oriented pre-tests in late February, and this was accordingly done. A standardized reading achievement test, a uniformly assigned written composition, and a reading attitude inventory were administered to the entire population. A subsample of each classroom was given an individual oral reading test, and a random selection of classrooms participated in an observation of reading as a free-time activity. Post-tests were administered in mid-May after approximately 165 days of reading instruction. They included alternate forms of the group and individual reading tests, a second sample of written composition, and an experimental "linguistic reading" achievement test, presently available in only one form.

Teacher-Training and Program Fidelity Control

Several investigators (e.g., Chall, 1963; Lerner, 1967) report a wide divergence in what a reading program's authors intend it to be and the way it is implemented in classrooms. Downing (1965) has pointed to setbacks to i/t/a instruction which result from faulty implementation. It was therefore extremely important to the validity of this research that there be close control of both methodology and materials in our experimental groups. Teachers could not simply be assigned materials and manuals and set forth upon the year of instruction. Similar concern for program implementation within the control groups was not deemed necessary. The control group was large, and use of varied materials and implementations was intended in establishing a norm for the district.

The majority of the experimental programs were assigned to teachers who had had no previous experience with the materials. None of them had had more than one year of (largely undirected) previous experience with their program. Therefore, early workshop instruction of the experimental teachers was initiated before school began, or early in the school year, as a first step toward control of fidelity to their programs. Whenever possible, specialists representing the publisher as authors, editors, or consultants worked with the teachers to develop an understanding of program use and to help solve possible problems. Released time for the teachers was provided by the district when workshop instruction was arranged during all or part of a regular school day. When publishers were unable to provide adequate training representatives, research staff or district personnel who were thoroughly familiar with the program(s) provided initial training which was augmented by conference telephone conversations with authors and/or consultants.

In addition to these early workshops, members of the research staff who had familiarized themselves with a particular program or programs held monthly meetings with their teachers to discuss program adherence and current problems. As a training tool to be used in these meetings and to aid in assuring fidelity, rating scales were developed by the staff member responsible for each program (see Appendix A). They were used during unscheduled monthly classroom observations of one-half hour on two consecutive days. While the teacher was not given advance warning of the visit, the principal was consulted to avoid days when school activities might conflict with classroom reading lessons. The control groups were also observed on an identical schedule, using a rating scale which was developed to indicate procedures within these groups which might otherwise have been thought to be the sole domain of an experimental group. There were no workshops scheduled for control group teachers, although this should be given consideration in future research.

No specific format was established for the development of the rating scales, though similarities do exist as a result of staff proximity and conferences during their development. The only guidelines were provided by their purpose as a training tool, usable by their developer, and a need to provide knowledge about the implementation of the most pertinent techniques and philosophies of the program. They are basically outlines of the major areas of each program, with allowance for indicating, through circled numbers and comments or yes-no checks, whether program adherence was lacking, only partial, or unabridged. The majority consist of a single sheet for recording, with directions for their use provided separately when the developer felt this was necessary. Copies of these scales have been provided in Appendix A, with proper recognition given to their authors.

While it was considered impractical to control for equal amounts of reading instructional time within classrooms, this was felt to be an important variable which could be reflected in the measured research results (Harris, 1967). A computer scorable time log (see Appendix B) was developed as a tool for accounting for the amount of teacher-directed reading instructional time in each classroom. To help assure uniformity, directions and criteria for the use of the log were discussed during a special after-school meeting of all participating teachers. For the months of November and December, while awaiting delivery of the printed forms, the teachers in all groups were asked to devote a minimum of one hour a day to reading instruction. For the full month of January and one week per month through April, they kept a daily log of time spent directing specified reading activities. Based on this sampling, mean teacher-directed instructional time served as a dependent variable for intercorrelation with achievement results.

In developing the time log sheet, it was necessary to first designate areas of classroom activity which contributed to all levels of the language program. Those directly related to reading instruction were delineated by the skills being taught as well as time devoted to instruction by ability groups (e.g., Group I was to be composed of the highest reading-grade children). Instructional time which was indirectly related to reading, such as listening or writing activities, was also to be designated. Since student-time would be difficult to assess for any activity, only the time in which the teacher was directly involved in a teaching activity was to be recorded. Design of the actual sheet was the responsibility of Mr. Ronald Beckstrom, district computer specialist. He planned for maximum information on a single sheet and was able to overcome a problem of interference of ink density, experienced by Harris (1967), through use of red ink.

While no attempt was made to control materials in other curriculum areas, supplemental reading materials used in the "linguistic" classrooms were controlled, on publisher recommendation, to prevent interference with the children's realization of the regularities of our written language. Experience charts which use a full range of vocabulary were discouraged. A bibliography of supplementary books which utilize regular sound-spelling correspondence was constructed after examination of books available in the district and was distributed to the teachers (see Appendix C).

General Administrative and Testing Procedures

This entire project was accomplished without full-time involvement by any member of the staff. A staff of thirty, including this author as director, consisted in large part of district personnel who were otherwise assigned but were released to devote time to the study as needed for classroom observations and/or testing. Graduate students from the University of Utah, one of whom was interested in portions of the data for a thesis, were also available for a few hours a week. A staff of four scorers, separate from the testing staff, was employed on an hourly basis and was trained in correcting each test by a staff member. Two consultant specialists participated in implementation of specific aspects of the design, principally in computer-related areas. In addition, a clerical staff of four transferred data to computer sheets used in key punching.

In consultation with Dr. Ethna R. Reid, Reading Center Director, and Dr. Gabriel Della-Piana, University of Utah Director of Educational Research and consultant to the Reading Center, the author, as project director, assumed responsibility for overall design of the research as well as the selection, and any necessary construction, of the testing instruments. Coordination of all activities, compilation of data, intra-study correspondence and the training of test personnel were also the responsibility of the project director.

Members of the staff were trained in test administration in order to assure uniformity and conformance to the procedures provided with each test instrument. Manuals of standardized tests and directions accompanying all others were rigidly adhered to, and necessary timing was done by stop watch. For group tests, classroom teachers served as proctors only to the extent of helping students keep their place or listening to the reading of written compositions. All directions were given by the test administrator. All individual testing was done in a quiet, private room other than the classroom. A statistical table of random numbers was used to select a subsample of children for these tests.

Through scheduling, time of day for testing was equalized among the programs, including the control groups. Enough testing personnel were used to complete a test series in a brief time that would avoid differences in the amount of teaching received by any one group prior to testing. All scoring was done in an agreed manner, following specific written directions (see Appendix D), by trained members of the scoring staff.

Instruments and Schedule of Testing

1. *Pre-tests.* The following reading-readiness and intelligence tests were administered to the entire population as pre-tests during the second and third week of the school year.

The Murphy-Durrell Reading Readiness Analysis (Murphy and Durrell, 1965)

Phonemes Test
Letter Names Test
Learning Rate Test

The Lorge-Thorndike Intelligence Test, Level I Form A (Lorge and Thorndike, 1957)

The *Murphy-Durrell Reading Readiness Analysis* was administered to the entire research population during the second week of the school year. Three subtests are included in the battery. The first, the Phonemes Test, is designed to measure a child's ability to identify separate sounds in words.

Initial consonant sounds, in the order of frequency in which they occur in the language, are tested first, followed by five sounds in final position. The phonemes are first taught and then tested by the placing of an X on pictures beginning or ending with the specified sound. The Letter Names Test contains 52 items which test a child's ability to identify letters, named by the test administrator, from a series of letters. Letters such as b, d, p, and q are tested against each other to assure that they are really distinguished. The Learning Rate Test is intended to test the number of sight words a child is capable of learning, following typical formal instruction procedures. Nine familiar words are taught, through use of chalkboard, flash cards, and test booklet, for identification in a multiple-choice test one hour later. The children must first discriminate between these and other words taught and then between these and untaught words of similar form.

The Cooperative First Grade Study (Bond and Dykstra, 1967) reported that the Letter Names subtest of the Murphy-Durrell was the best single predictor of first grade reading achievement. The Phonemes Test ranked next in value as a predictor.

The *Lorge-Thorndike Intelligence Test* was administered as a group test during the third week of the school year. It is an untimed, power test of abstract intelligence containing only pictorial items. Three subtests include oral vocabulary, selection of inappropriate items, and selection of paired pictures. The instructions are oral, repeated and reinforced with each item, and are paced to the group by the administrator.

2. *The mid-year tests.* The mid-year series of tests, administered in February and early March, served in part as skill-oriented pre-tests for the study and included the following instruments.

The Gates-MacGinitie Primary A (B) Reading Test, Form I (Gates and MacGinitie, 1966)

Vocabulary
Comprehension

The Gilmore Oral Reading Test, Form A (Gilmore, 1951)

Accuracy
Comprehension
Rate

Sample Written Composition

Number of words attempted
Subordinate Clause Index
Mean T-Unit Length

The San Diego Inventory of Reading Attitudes (Vogt, Lane, Kellogg, and Norin, 1964)

Observation of Motivation to Read

The Gates-MacGinitie Reading Test is part of a new series which replaces the *Gates Primary Reading Test.* It consists of two parts and is available in two equivalent forms, I and II. The first part, a Vocabulary Test, consists of exercises which contain four words and a picture illustrating the meaning of the word to be selected. The Comprehension Test consists of items containing a panel of four pictures from which to select the one that illustrates accompanying paragraphs of increasing length and difficulty. Although it is not a speed test, a time limit is specified for each subsection. Children who tested within one point of the top in either vocabulary or comprehension were retested with the next level Primary B (Form I) test which has a higher range.

The *Gilmore Oral Reading Test* was administered to a randomly selected subsample of five children in each classroom. It consists of a series of graded paragraphs to be read aloud following discussion of a motivating picture. Errors are recorded by type (e.g., omissions, substitutions, etc.) and comprehension is checked through questions provided with each paragraph. In addition to the accuracy, comprehension, and rate scores of this test, the number and kind of errors made could be computed and used as dependent variables. It may be possible that word attack differentials between programs will appear. For instance, significantly more frequent substitution or mispronunciation errors by children instructed by one method may indicate a greater willingness to attack unfamiliar words than shown by a program in which the children predominantly required tester prompting of unfamiliar words.

A written composition was assigned to the entire population by research staff members during this test series. Since it is an area in which there is considerable current interest but little precedent, particularly on the first grade level, techniques for assessing control of written expression were developed for this project. A dialogue, including a motivational poem (see Appendix E), was provided to hold procedures consistent in all classrooms. A poem was selected for subject motivation for reasons of brevity and controllability. (A story would involve more time and subject-oriented classroom motivational conversation could not be entirely uniform in all classrooms, even with the use of identical stimulus pictures or objects.) The children were asked to write their own stories, making them as long as they liked, and making their best attempt to spell each word themselves. No help in spelling was given by the administrator. At the end of twenty minutes, or earlier if they were finished, each child was asked to read his story to the staff member or teacher to assure proper interpretation of the words intended. Obscurely spelled words were written in red above the lines by the listener. Words, or inflections, which were read but had been omitted in the written form, were also added in red above the line and enclosed in parentheses, in order to facilitate interpretation of sentence structure.

In the past, evaluation of written composition has been almost entirely limited to computation of the numbers of words written, mean sentence length, and attempts, at best subjective, to determine creativity. The last-named remains an elusive and subjective judgment and, as such, was not an aspect of the evaluation used in this study. The weakness of relying upon sentence length as a developmental indicator, particularly in primary writing which is apt to contain "run-on" sentences extended by conjunctions, has recently been recognized. Many of the alternatives to the use of this index have been summarized by Cartwright (1966). Hunt (1965; 1966) has developed extremely influential measures of syntactic development which are based upon linguistic transformational theory. In large measure, past research has, like Hunt's, been limited to the upper elementary grades and above. For this study, it was felt that it would be valuable to selectively adapt these previous techniques to our first-grade population.

Three measures were determined to be most significant for our population. One of them, number of words attempted, was judged to be the foundation upon which higher order syntactic control would be built. A "subordinate clause index," as modified from the original by Hunt (1966), was obtained by the following formula:

$$\frac{\text{Total No. of main clauses} + \text{Total No. of subordinate clauses}}{\text{Total No. of main clauses}} =$$

The Subordinate Clause Index

Hunt found that an increase in this index consistently reflects the general trend of language development. It would also reflect the most important difference found by Fries (1940) in the span from semiliterate to standard written English of adults.

The "minimal terminal units" (T-Units) of Hunt (1966) were utilized in the remaining measurements through use of the following formula:

$$\frac{\text{Total No. words attempted}}{\text{Total No. of T-Units}} = \text{Mean T-Unit Length}$$

A T-unit is definable as the "shortest unit into which a piece of discourse can be cut without leaving sentence fragments as residue" (Hunt, 1966, p. 737). In size, they intervene between the main clause and the compound sentence, since each one consists of a main clause plus whatever subordinate clause(s) is attached to or embedded within it. The measurement of mean T-unit length has been shown to reflect a developmental tendency to make use of more subordinate clauses per main clause as writing becomes more mature and controlled. Griffin (1965) further utilized the T-unit to compute an index of the number of generalizing transformations performed in written sentence construction. This was, however, felt to be an unnecessary measure in evaluation of primary writing samples.

Anderson (1937), in testing the reliability of several clausal measures, found them to be unreliable when two or more scorers worked on the same set of compositions. In the present study, therefore, one scorer, a doctoral candidate in linguistics who was thoroughly familiar with the terminology, was used to score the entire group of compositions. A criterion sheet (see Appendix F) for use in scoring was supplied to provide for adherence to the intentions of the researcher. Since the children instructed in i/t/a were in various stages of transfer to TO, an i/t/a alphabet key was provided the scorer and a cross-check of the phonetic acceptability of the spellings was made with the classroom teacher as well.

To conclude this series of mid-tests, attitude toward reading and an inclination (motivation) to read outside the instructional situation were both deemed important aspects of successful reading instruction. A standardized *Inventory of Reading Attitude,* developed by the San Diego County Public Schools, was administered, with the permission of that body, to the entire population early in March. Statements concerning reading were read to the children by the test administrator. They were told to circle either yes or no on the inventory sheet to show how they usually felt about reading. The 25 items included in the inventory were selected as the most discriminatory from an original list of 114 items administered in 24 classrooms. The first three questions of the inventory are included below to exemplify the design of this instrument.

Yes No 1. Do you like to read before you go to bed?
Yes No 2. Do you think that you are a poor reader?
Yes No 3. Are you interested in what other people read?

A random sampling of classrooms participated in a pilot study intended to assess motivation to read outside the instructional situation. Each of these classrooms was given ten minutes of "free time" twice a week over a three-week period (see Appendix G). One of the periods was to occur early in the week, morning or afternoon, and the other was to be late in the week, at an opposite time of day. The children were given a choice of art, writing, or reading activity and were told they must stay with their choice for the entire ten-minute period. The teacher noted on the scoring sheet, with checks, all children who selected the reading activity as their choice. To provide a scoring range, the six possible selections of reading per child were weighted by allowing one point for the first reading choice and an additional two points for each subsequent choice of reading (e.g., one reading selection = one point, two reading selections = three points, etc. to a possible total of eleven points).

3. *Post-tests.* Post-tests were administered in mid-May after approximately 165 days of instruction, and included the following instruments.

The Gates-MacGinitie Primary A Reading Test, Form II (Gates and MacGinitie, 1966)

Vocabulary
Comprehension

The Gilmore Oral Reading Test, Form B (Gilmore, 1951)

Accuracy
Comprehension
Rate

Linguistic Reading Test, Grade 1 (Anastasiow and Hansen, 1966)

Sample Written Composition

Number of words attempted
Subordinate Clause Index
Mean T-Unit Length

The alternate form of the previously discussed Gates-MacGinitie test was administered in only the A level of difficulty. A problem in comparing the February scores of the two levels resulted in the omission of the B level in this test series. The written composition was assigned in a manner comparable to the pre-test. The individual *Gilmore Oral Reading Test* was administered to the same subsample tested in February.

The *Linguistic Reading Test* is an experimental test which was given to a random sampling of "linguistic" and control classrooms. It was developed for the Palo Alto Unified School District to supplant traditional reading measures which are recognized to handicap the "linguistically" trained readers during the early stages of instruction. Varied aspects of the decoding process are measured by this test as pupils progress through seven subtests. For instance, Test 2 measures knowledge of grapheme-phoneme correspondence rules by requiring the student to make changes such as "not" to "note." Test 5 requires the selection of grammatical phrases such as "by the pond" rather than "pond the by." The seventh test necessitates listening to utterances such as "I gotta" and associating them with their written form, "I've got to." Use thus far by the developers seems to indicate that the test is a reliable measure of reading ability in conventional as well as "linguistic" programs (Anastasiow and Hansen, 1967).

Simply for reasons of exploratory interest to future research rather than as a source of immediate significance as data, a small subsample of children was used to provide short samples of tape-recorded oral conversation related in subject to the written compositions. These were evaluated using the syntactic measures, and scorer, employed for the written composition. The purpose of this sample was to investigate the possibility of assessing differences among programs designed to produce proximity of written control of language to oral language control.

Statistical Treatment of Data

Statistical procedures were selected which were felt to be most efficient in providing information relevant to the purposes of the research design. Twenty-one dependent variables resulted from the data collected in the study. Within each program, excepting i/t/a, the Pearson product-moment correlation coefficient was utilized to determine the relationships between pupil/class characteristics, readiness, and achievement measures, and six 21 X 21 matrixes resulted.

Another important anticipated outcome of the research design was determination of whether different approaches to initial reading instruction would produce differentiated results in terms of the measured level of intelligence or readiness at time of the pre-tests. Therefore, the pupils within each program were blocked in turn according to their performance on each of these measures. Comparisons among the seven programs were performed for high (110 and above), middle (90 to 109), and low (89 and below) IQ children. A 3 (levels of IQ) X 7 (reading programs) analysis of variance was performed for each of the twenty-one dependent variables resulting from the study. Additionally, a 3 (levels of readiness) X 7 (reading programs) analysis of variance was performed on the same twenty-one dependent variables. To obtain the blocking for the three Murphy-Durrell readiness levels, all total test scores were tabulated and divided into three ranges of essentially equal population proportion throughout the programs. The high group included scores from 98 to 118, the middle group 74 to 97, and the low group 16 to 73.

The computations necessary for the three-level analysis of variance were performed by a Univac 1108 Computer using a two-way analysis of variance for unequal *n's* program as written by Richard Maxwell of the University of Utah Computer Center. Intercorrelations were computed using options from the Factor Analysis Program as written by Dan Anderson and Ed Dallin of the Computer Center. Michael Hogben was consultant on design and coordinated all data processing from initial direction of data coding through final computer analysis and reporting format.

CHAPTER IV

STATISTICAL RESULTS

The preliminary nature of this study and the imbalance in the number of subjects within the participating programs greatly restricts desirable statistical analysis and interprogram comparisons. Therefore, the following discussion of results will be limited to summary statements and concise tables. The complete data resulting from the study is available as a separate publication from the ECRI offices of Granite School District.

Relationships Between Pre-test Measures and Reading Achievement

The Pearson product-moment correlation coefficient was utilized to determine the relationships between pre-test readiness for reading and post-test reading achievement within each separate program. The six programs include a control group, instructed with traditional district basal reading texts (American Book, Allyn & Bacon, Ginn, Houghton-Mifflin, or Macmillan), and experimental groups instructed with the McGraw-Hill *Programmed Reading Series,* the *SRA Basic Reading Series,* the Harper & Row *Linguistic Readers,* the *Listen, Look, Learn* (LLL) materials, or the Educational Developmental Laboratories (EDL) machines as supplement to one of the district basal series.[1] The five readiness measures obtained were the Murphy-Durrell Phoneme, Letter Name, Learning Rate, and Total Test scores and the DIQ of the *Lorge-Thorndike Intelligence Test.* The Vocabulary and Comprehension subtests of the *Gates-MacGinitie Primary Reading Test,* Form I, administered mid-year, served as a pre-test measure of reading ability. Form II of the Gates-MacGinitie Test, administered at the end of the experimental period, was selected as the prime achievement measure. The within-program means for each of the dependent variables of the study are reported in Table 1. Due to sampling procedures in the administration of the *Gilmore Oral Reading Test,* a sufficient *n* was available for only the Control-McGraw-Hill comparison. Thus, the Gilmore results have been excluded from this report and will appear in a separate report.

A summary of the relationships of the pre-test readiness and reading measures to the Gates-MacGinitie Vocabulary achievement measure is re-

---

[1] The data from the single i/t/a classroom was eliminated from the reported statistics because of a severe imbalance and top-heaviness when blocking by readiness and IQ. A duplicate multiple-class i/t/a study was undertaken in Granite School District during the 1967-68 school year and data from that study is available as a separate publication from ECRI.

Table 1 Within-program Means and Standard Deviations for each of the dependent variables of the study

	Control		McGraw-Hill		EDL-Basal		Listen Look Learn		SRA		Harper & Row	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Murphy-Durrell Phonemes	36.56	8.85	34.53	9.18	39.63	7.54	33.85	9.06	32.67	10.25	36.04	8.23
Murphy-Durrell Letter Names	37.78	12.11	38.15	12.32	41.53	10.70	31.42	14.53	31.42	14.53	39.32	10.70
Murphy-Durrell Learning Rate	10.65	5.01	8.93	6.40	12.16	3.77	9.47	4.22	8.46	4.71	9.46	4.36
Murphy-Durrell Total Score	84.93	21.06	81.37	20.49	93.03	19.65	83.50	20.72	72.19	16.72	84.11	19.02
Lorge-Thorndike DIQ	101.91	12.57	97.57	12.46	99.14	11.91	100.88	11.74	98.00	14.67	101.93	11.49
Sex	0.53	0.50	0.50	0.50	0.50	0.50	0.44	0.50	0.48	0.50	0.57	0.50
Gates-MacGinitie I Vocabulary	27.70	9.89	29.21	11.02	25.45	10.00	26.54	10.25	26.00	13.00	32.04	7.99
Gates-MacGinitie I Comprehension	15.12	7.93	16.06	8.82	13.22	6.47	15.67	5.93	15.23	7.62	18.88	5.64
Gates-MacGinitie II Vocabulary	36.41	10.07	37.77	9.72	37.13	9.08	35.78	10.34	30.85	12.99	41.92	4.15
Gates-MacGinitie II Comprehension	21.65	7.53	23.24	8.42	21.60	7.68	21.86	7.41	18.02	9.29	24.38	5.96
Attitude toward Reading	5.40	1.71	5.51	1.58	5.05	1.90	5.52	1.87	4.83	2.28	4.59	1.60

Table 1 Within-program Means and Standard Deviations for each of the dependent variables of the study

	Control		McGraw-Hill		EDL-Basal		Listen Look Learn		SRA		Harper & Row	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Motivation to Read	4.88	3.60	4.03	3.68	4.77	3.40	0.96	1.60	1.41	1.43	no data	
Written Sample I No. of Words	23.04	13.38	21.00	14.23	22.07	12.97	16.61	9.09	18.71	15.06	21.67	7.55
Written Sample I Clause Index	1.16	0.77	1.23	1.01	0.94	0.52	1.04	0.47	0.91	0.59	1.23	0.34
Written Sample I T-unit length	6.37	5.27	5.79	3.33	5.97	4.80	6.88	4.19	5.67	6.05	6.33	2.61
Written Sample II No. of Words	26.92	15.42	33.64	19.81	34.74	22.96	27.38	12.97	16.63	10.40	24.00	5.91
Written Sample II Clause Index	1.28	0.55	1.35	0.69	1.31	0.69	1.50	0.73	0.85	0.67	1.80	0.82
Written Sample II T-unit length	9.05	5.56	8.38	4.59	8.27	4.67	9.52	4.56	4.67	3.61	10.62	3.27
Socio-Economic	3.26	1.19	3.68	1.05	3.93	1.19	3.86	1.80	4.00	0.00	2.00	0.00
Linguistic Reading Test	35.00	13.49	41.92	13.54	36.14	15.09	no data		34.10	14.31	44.35	12.36
Total Reading Time	165.61	50.12	127.20	49.08	161.85	35.04	148.79	39.31	99.10	43.15	131.20	0.04

S.D. = Standard Deviation

Table 2 Correlations between pre-test measures and the post-test Vocabulary subtest of the Gates-MacGinitie Primary Reading Test, Form II

	Control	McGraw-Hill	EDL-Basal	Listen Look Learn	SRA	Harper-Row
Mean Ns	404	268	166	88	39	23
Murphy-Durrell Phonemes	.49	.51	.66	.33	.64	.11*
Murphy-Durrell Letter Names	.59	.56	.69	.67	.76	.54
Murphy-Durrell Learning Rate	.36	.27	.55	.53	.45	.37*
Murphy-Durrell Total Score	.65	.65	.75	.63	.74	.52
Lorge-Thorndike Intelligence Test	.35	.44	.37	.29	.69	.38*
Gates-MacGinitie I Vocabulary	.73	.76	.75	.80	.90	.11*

*Did not reach the .05 level of significance. All other correlations reached the .01 level of significance.

ported in Table 2. In recognition of the interprogram numerical imbalance in subjects, the mean number of subjects contained in intraprogram correlations has been included in some tables. The Murphy-Durrell Total Test score was the best predictor of achievement in three of the groups and a strong second in predictive rank in the other three. All of these correlations were substantial, ranging from .52 to .75. However, the Letter Names subtest, with a correlation range of .54 to .76, was the most important predictive component of the Murphy-Durrell in all programs. The Cooperative First Grade Study (Bond and Dykstra, 1967) also found this highly significant relationship between pre-instruction ability to recognize letters and success in learning to read in first grade. In four of the programs, the Phonemes subtest, which assesses ability to differentiate sounds in words, ranked second as a predictive component of the Murphy-Durrell. This included programs which, because they employ the traditional basal teaching methods utilized in the Learning Rate subtest, might be expected to have a higher achievement relationship to that subtest. In no reading program was the relationship of the Lorge-Thorndike DIQ to achievement as high as that of either the Letter Names subtest or Total Test score of the Murphy-Durrell. The high correlations of the Gates-MacGinitie pre-test and post-test Vocabulary scores are predictable for two parallel forms of a test.

Mention should be made here of the unconformity of the Harper-Row correlations throughout the data. A proportionate balance of subjects in this classroom across readiness and IQ levels permitted the inclusion of these results. However, it is felt that the use of a single classroom unit and the resultant small *n* invalidates any conclusions which might be drawn from these correlations. It is not possible to isolate the factor(s) which might have been operating in that one classroom.

The relationship of each of the pre-test measures to the Comprehension subtest of the Gates-MacGinitie Test is summarized in Table 3. The Murphy-Durrell Total Test score, with correlations ranging from .60 to .72, was the best predictor of achievement in every program except Harper-Row. Ranking first in that program, and being a strong second in all others, the Letter Names subtest, with a range of .52 to .79, was again the best predictive component. The Lorge-Thorndike DIQ, with a range of .31 to .66, was lower than either the Letter Names or Total Test correlations in five programs and equal to the second-ranking Letter Names relationship in the SRA program. The correlations of the Gates-MacGinitie pre-test and post-test Comprehension scores, ranging from .45 to .71, are somewhat lower than those of the Vocabulary scores. They are thought to

Table 3 Correlations between pre-test measures and the post-test Comprehension subtest of the Gates-MacGinitie Primary Reading Test, Form II

	Control	McGraw-Hill	EDL-Basal	Listen Look Learn	SRA	Harper-Row
Mean Ns	404	268	166	88	39	23
Murphy-Durrell Phonemes	.49	.43	.59	.35	.62	.25*
Murphy-Durrell Letter Names	.55	.52	.66	.57	.66	.79
Murphy-Durrell Learning Rate	.37	.31	.61	.49	.42	.34*
Murphy-Durrell Total Score	.63	.60	.72	.60	.67	.64
Lorge-Thorndike Intelligence Test	.34	.45	.38	.31	.66	.39*
Gates-MacGinitie I Comprehension	.45	.58	.67	.62	.71	.61

*Did not reach the .05 level of significance. All other correlations reached the .01 level of significance.

Table 4 Correlations between pupil/class characteristics, written composition samples, Linguistic Reading Test, and the Murphy-Durrell Total Test scores

	Control	McGraw-Hill	EDL-Basal	Listen Look Learn	SRA	Harper-Row
Sex	-.15	-.12	-.09	-.06	-.10	-.18
Socio-Economic	.08	-.20	-.18	-.28	.00	.00
Attitude toward Reading	.17*	.17*	.24*	.07	.02	-.43
Motivation to Read	.06	.17	.24	.18	.44*	
Written Sample I No. of Words	.39*	.34*	.39*	.34*	.31	.12
Written Sample I Clause Index	.16*	.20*	.32*	.20	.35	.11
Written Sample I T-unit length	.16*	.20*	.27*	.05	.41*	.08
Written Sample II No. of Words	.28*	.03	.42*	.15	.45	.30
Written Sample II Clause Index	.18	.08	.28	.31	.12	.12
Written Sample II T-unit length	.28*	.08	.36*	.19	.26	.62
Linguistic Reading Test	.60*	.54*	.83*		.66*	.78*
Lorge-Thorndike Intelligence Test	42*	.45*	.44*	.44*	.68*	.49*

*.01 level of significance.

reflect a trend toward reading for comprehension taking place during the latter half of first grade.

Relationships of Pre-test and Achievement Measures to Pupil/Class Characteristics and Secondary Achievement Scores

The correlations between pupil/class characteristics, written composition samples, the *Linguistic Reading Test* and the Murphy-Durrell Total Test readiness score are summarized in Table 4. Because the intra-program range of *n* for these measures is large, due to sampling procedures, the table indicates those correlations reaching the .01 level of significance rather than reporting the mean *n* for each program. In general, low correlations with the variables of sex (-.18 to -.06), socio-economic status (-.28 to .08), attitude toward reading (.02 to .24), and motivation to read (.06 to .44), indicate little relationship of any of these variables to a child's readiness level. DIQ, with correlations ranging from .44 to .68, was the most important contributing factor in readiness performance. Although perhaps related to reading achievement, the measures of unaided written composition, obtained both mid-year and as post-tests, resulted in lower relationships to pre-instructional readiness than did the reading achievement test results. The added measurement of facility with oral language in future research may provide insight into related skills missed by a readiness test. The high correlations of the unorthodox *Linguistic Reading Test* with readiness, ranging from .54 to .83, were similar to those of the more traditional Gates-MacGinitie test and were not selective of the "linguistically" trained groups.

The relationships of the subtests of the Gates-MacGinitie test to pupil/class characteristics and secondary achievement scores are summarized in Tables 5 and 6. The variables of sex, socio-economic status, attitude toward reading, and motivation to read as a free-time activity have a generally low correlation with measured reading achievement. However, statistically significant correlations did result for the attitude and motivation variables in the control groups and McGraw-Hill program, which would have practical value only for the extremes in these groups. The variable of average reading-instructional time within each classroom bore little or no relationship to the achievement scores. The relationships of the measures of written language to achievement range from moderate to low. They are apt to be highest in the number of words written rather than in the complexity or length of the constructions. Perhaps account needs to be taken of oral language construction complexity and length as a guideline against which to measure written language. Very high corre-

Table 5 Correlations between pupil/class characteristics, written composition samples, Linguistic Reading Test, average reading-instructional time, and the Gates-MacGinitie, Form II, Vocabulary scores

	Control	McGraw-Hill	EDL-Basal	Listen Look Learn	SRA	Harper-Row
Sex	-.05	-.14	-.13	-.11	-.26	.10
Socio-Economic	.03	-.24	-.13	-.09	.00	.00
Attitude toward Reading	.16*	.16	.25*	.18	.32	-.02
Motivation to Read	.28*	.32*	.25	.19	.39*	
Written Sample I No. of Words	.39*	.36*	.40*	.39*	.46*	.20
Written Sample I Clause Index	.22*	.19*	[illegible]	.18	.59*	.08
Written Sample I T-unit length	.20*	.24*	.24*	.05	.41*	.02
Written Sample II No. of Words	.36*	.20	.42*	.29	.35	.17
Written Sample II Clause Index	.16	.21	.37*	.26	.36	.04
Written Sample II T-unit length	.22*	.32*	.37*	.22	.49	.35
Linguistic Reading Test	.72*	.78*	.85*		.84*	.77*
Average Instructional Time	.17*	.15	.07	.01	.19	.00

*.01 level of significance.

Table 6 Correlations between pupil/class characteristics, written composition samples, Linguistic Reading Test, average reading-instructional time, and the Gates-MacGinitie, Form II, Comprehension scores

	Control	McGraw-Hill	EDL-Basal	Listen Look Learn	SRA	Harper-Row
Sex	-.14	-.17	-.10	-.06	-.31	-.17
Socio-Economic	.03	-.19	-.14	-.16	.00	.00
Attitude toward Reading	.21*	.19*	.29*	.19	.37	.03
Motivation to Read	.32*	.37*	.26	.17	.35	
Written Sample I No. of Words	.42*	.32*	.46*	.40*	.34	.63*
Written Sample I Clause Index	.21*	.15	.35*	.12	.40*	-.08
Written Sample I T-unit length	.22*	.15	.22*	-.04	.35	-.10
Written Sample II No. of Words	.30*	.17	.46*	.29	.48	.42
Written Sample II Clause Index	.07	.24	.36*	.22	.32	.00
Written Sample II T-unit length	.22*	.20	.44*	.21	.46	.67
Linguistic Reading Test	.84*	.46*	.85*		.82*	.82*
Average Instructional Time	.18*	.16*	.03	.01	.07	.00

*.01 level of significance

lations between the *Linguistic Reading Test* and the two traditional achievement scores indicate that they measure very similar characteristics even though their test form and approach is quite different. With the exception of a McGraw-Hill correlation of .46 between the Comprehension score and the Linguistic test, the two tests appear to test all programs without prejudice, making the use of two tests seem unnecessary.

Evaluation of Comparative Program Effectiveness by Readiness Level

An important anticipated outcome of this study was the determination of whether any of the different approaches to initial reading instruction would produce differentiated results in terms of the measured level of readiness or intelligence at time of pre-tests. Pupils were blocked in turn according to their performance on the Murphy-Durrell total test and again according to their Lorge-Thorndike DIQ. A 3 (levels of readiness) X 6 (Program) analysis of variance was performed for each of the dependent variables for high (98 to 118), middle (74 to 97), and low (16 to 73) readiness total test scores. Additionally, a 3 (levels of IQ) X 6 (Program) analysis of variance was performed for the dependent variables for high (110 and above), middle (90 to 109), and low (89 and below) IQ children. To obtain the blocking, all total test readiness scores and DIQs were tabulated and divided into ranges of approximately equal population proportion throughout the programs. Again, the extreme variation in *n* among the programs resulted in a proportionate variation among programs at each level and is a limiting factor in the discussion of significant results. Only those programs with comparable *n*'s can be directly contrasted with any confidence. All other contrasts must be against the control group, established as a district norm.

Effectiveness blocking on Murphy-Durrell performance. In general, dependent variables which proved to have a low relationship to achievement are not included in this discussion. However, the relationships of socio-economic status to achievement, discussed in a previous section, were obscured by the fact that both the extreme high socio-economic program (Harper-Row) and the extreme low (SRA), as ranked by intra-program across-levels means, were homogeneous, one-school programs with S.D.s of 0.00. All other programs, with greater spans of classrooms and schools, were more heterogeneous. Since, in distributing the students of these programs to levels according to pre-test performance, it was possible for intraprogram subjects with a common socio-economic value to

Table 7 Socio-economic ranking of programs by unadjusted, across-levels means of measured socio-economic values

Program	Rank*	N	Mean	S.D.
Harper-Row	1	28	2.00	0.00
Control	2	493	3.26	1.19
McGraw-Hill	3	323	3.68	1.05
Listen, Look, Learn	4	103	3.86	1.80
EDL-Basal	5	202	3.93	1.19
SRA	6	48	4.00	0.00

*1 = the group highest in socio-economic status.

Table 8 F-ratios for reading achievement measures blocking on levels of Murphy-Durrell Total Test readiness

	Gates-MacGinitie I (Mid-test)		Gates-MacGinitie II (Post-test)		
	Vocabulary	Comprehension	Vocabulary	Comprehension	Linguistic Reading Test
Reading Programs (RP)	7.11*	5.25*	8.07*	6.37*	3.97**
Levels (L)	323.00*	98.00*	285.00*	257.00*	47.60*
RP x L	9.36*	7.20*	3.04**	4.43*	1.36

*.001 level of significance
**.01 level of significance

cluster in one level, close examination of the data was indicated. Both significant program differences and program-by-level interactions resulted in the Murphy-Durrell readiness-levels analysis of variance. The data revealed several significant interprogram differences in readiness levels, many of which acted to change the socio-economic ranking severely. As a frame of reference, Table 7 provides a socio-economic ranking of the programs by their across-levels unadjusted means of that value. Programs with the lowest means have the highest socio-economic ranking, since the values were computed on a range from 1 (highest) to 5 (lowest). Variations in this ranking, as well as variations due to clustering in sex (ratio of boys to girls) and IQ, were explored in the consideration of all significant program differences. They are indicated when found to be pertinent to these differences.

The results of the analysis of interprogram reading achievement blocking on Murphy-Durrell readiness levels are summarized in Table 8. Significant program differences and program by level interactions were found for the subtests of both forms of the Gates-MacGinitie, indicating that the programs operated differently for students of varying readiness. For each variable, fifteen interprogram comparisons were made for each level. Only a review of the resulting significant differences and their implications is included in this report. Results of interprogram T-Tests, blocking on levels of Murphy-Durrell Total Test readiness, are reported in Tables 8 through 23. More detailed information will be found in consulting the separate publication of the project statistical tables which is available through ECRI.

For the high readiness students, the significant differences favored McGraw-Hill *Programmed Reading* over the Control and EDL-Basal on all four Gates-MacGinitie measures and also over the *Listen, Look, Learn* (LLL) on the Form I vocabulary measure. The only other significant differences for this level favored SRA over EDL-Basal on the Form I subtests and the Controls over EDL-Basal on the Form I vocabulary test. In this level, significant differences for the *Linguistic Reading Test* also favored McGraw-Hill.

In the middle readiness level, McGraw-Hill, SRA, and Harper-Row were all generally significantly more effective than the Control, LLL, or EDL-Basal programs on all four Gates-MacGinitie measures. Results in this level for the Linguistic test favored SRA and Harper-Row over the other groups.

Within the lowest level of readiness, a different pattern resulted. For the Form I subtests, it was a matter of all other programs being signifi-

Table 9 T-Tests between the *Control* and *McGraw-Hill Programmed Reading* groups for selected dependent variables blocking on levels of Murphy-Durrell Total Test readiness

	Level I Total Score: 98-118		Level II Total Score: 74-97		Level III Total Score: 16-73	
Dependent Variable	t-value	S.D.	t-value	S.D.	t-value	S.D.
Sex	0.96	0.07	-0.16	0.06	1.54	0.06
Gates-MacGinitie I Vocabulary	-4.62*	0.99	-3.29*	1.01	-0.35	1.17
Gates-MacGinitie I Comprehension	-4.84*	0.99	-2.65*	0.83	1.86	1.26
Gates-MacGinitie II Vocabulary	-2.02***	0.76	-3.28*	0.97	-1.47	1.49
Gates-MacGinitie II Comprehension	-2.68**	0.85	-3.98*	0.87	-1.27	0.94
Attitude toward Reading	-0.46	0.26	-1.17	0.20	-0.62	0.23
Motivation to Read	0.55	0.75	1.51	0.66	1.38	0.71
Written Sample I No. of Words	1.44	2.07	-0.29	1.46	1.00	1.78
Written Sample I Clause Index	-2.52**	0.13	-0.15	0.11	0.13	0.11
Written Sample I T-unit length	-0.11	0.55	1.76	0.67	0.06	0.50
Written Sample II No. of Words	-0.98	3.69	-2.20***	2.94	-2.41**	5.83
Written Sample II Clause Index	0.48	0.09	-1.22	0.14	-0.60	0.13
Written Sample II T-unit length	1.86	1.07	0.02	1.08	-0.76	1.07
Lorge-Thorndike Intelligence Test	1.72	1.64	2.06***	1.33	3.26*	1.51
Socio-Economic	1.12	0.17	-4.20*	0.13	-6.18*	0.13
Linguistic Reading Test	-2.53**	3.52	-0.26	3.51	-3.09*	4.36
Total Reading Time	4.37*	7.22	7.15*	5.48	6.65*	5.99

Minus sign on t-value indicates that the second program has a higher mean
*.01 level of significance
**.02 level of significance
***.05 level of significance

Table 10 T-Tests between the *Control* and *SRA Basic Readers* groups for selected dependent variables blocking on levels of Murphy-Durrell Total Test readiness

	Level I Total Score: 98-118		Level II Total Score: 74-97		Level III Total Score: 16-73	
Dependent Variable	t-value	S.D.	t-value	S.D.	t-value	S.D.
Sex	0.65	0.17	0.47	0.14	0.98	0.10
Gates-MacGinitie I Vocabulary	-1.27	2.44	-3.39*	2.94	1.25	1.97
Gates-MacGinitie I Comprehension	-1.74	2.27	-3.42*	1.82	0.93	2.46
Gates-MacGinitie II Vocabulary	0.07	1.96	-1.38	2.85	2.01***	2.43
Gates-MacGinitie II Comprehension	-0.61	2.13	-0.31	2.30	2.00***	1.42
Attitude toward Reading	1.94	1.23	0.18	0.59	0.70	0.38
Motivation to Read	2.69*	1.35	2.46**	1.20	4.64*	0.70
Written Sample I No. of Words	1.27	4.97	-1.05	3.65	1.09	3.06
Written Sample I Clause Index	0.88	0.20	0.36	0.28	1.20	0.19
Written Sample I T-unit length	-2.20	1.57***	0.40	2.16	1.22	0.91
Written Sample II No. of Words	0.76	6.87	1.55	8.07	1.94	4.02
Written Sample II Clause Index	0.50	0.19	2.57**	0.32	1.20	0.23
Written Sample II T-unit length	0.91	2.62	2.23***	2.86	1.86	1.40
Lorge-Thorndike Intelligence Test	-1.27	4.03	-0.53	3.32	2.21***	2.62
Socio-Economic	-1.36	0.41	-2.32***	0.33	-4.23*	0.22
Linguistic Reading Test	-0.39	4.49	-2.25***	4.90	0.02	2.41
Total Reading Time	4.33*	17.77	4.17*	13.72	6.06*	10.41

Minus sign on t-value indicates that the second program has a higher mean
*.01 level of significance
**.02 level of significance
***.05 level of significance

Table 11 T-Tests between the *Control* and *Harper & Row Linguistic Readers* groups for selected dependent variables blocking on levels of Murphy-Durrell Total Test readiness

	Level I Total Score: 98-118		Level II Total Score: 74-97		Level III Total Score: 16-73	
Dependent Variable	t-value	S.D.	t-value	S.D.	t-value	S.D.
Sex	-1.07	0.21	0.46	0.13	-0.40	0.19
Gates-MacGinitie I Vocabulary	-0.80	3.57	-3.02*	2.19	-1.55	3.60
Gates-MacGinitie I Comprehension	-1.13	3.31	-4.80*	1.36	-0.06	4.69
Gates-MacGinitie II Vocabulary	-0.88	2.43	-2.05***	2.31	-2.85*	4.42
Gates-MacGinitie II Comprehension	0.15	2.70	-2.02*	1.87	-0.94	2.49
Attitude toward Reading	3.40*	0.70	1.71	0.49	-1.18	0.60
Motivation to Read	Insufficient Observations for Calculations					
Written Sample I No. of Words	1.80	5.61	-0.63	2.99	-0.52	5.09
Written Sample I Clause Index	-0.25	0.24	0.00	0.23	-0.65	0.32
Written Sample I T-unit length	0.39	1.54	0.35	1.84	-0.91	1.50
Written Sample II No. of Words	Insufficient Observations		-0.49	8.08	0.13	5.82
Written Sample II Clause Index	Insufficient Observations		-2.44**	0.33	-1.71	0.32
Written Sample II T-unit length	Insufficient Observations		-1.02	2.85	-0.86	1.96
Lorge-Thorndike Intelligence Test	-0.26	4.82	-0.30	3.21	0.22	4.78
Socio-Economic	2.85*	0.51	4.00*	0.31	2.60**	0.42
Linguistic Reading Test	-1.59	5.23	-3.68*	3.72	-0.98	4.03
Total Reading Time	1.86	21.39	2.77*	12.46	1.45	18.82

Minus sign on t-value indicates that the second program has a higher mean

*.01 level of significance

**.02 level of significance

***.05 level of significance

Table 12 T-Tests between the *Control* and *Listen, Look, Learn* groups for selected dependent variables blocking on levels of Murphy-Durrell Total Test readiness

	Level I Total Score: 98-118		Level II Total Score: 74-97		Level III Total Score: 16-73	
Dependent Variable	t-value	S.D.	t-value	S.D.	t-value	S.D.
Sex	0.66	0.10	0.48	0.08	2.11***	0.10
Gates-MacGinitie I Vocabulary	-0.05	1.58	1.09	1.41	0.36	2.05
Gates-MacGinitie I Comprehension	-1.18	1.43	-0.61	0.85	0.00	2.53
Gates-MacGinitie II Vocabulary	-1.03	1.09	0.43	1.51	0.41	2.32
Gates-MacGinitie II Comprehension	-0.98	1.21	-0.49	1.19	0.17	1.32
Attitude toward Reading	0.00	0.37	-0.18	0.32	-0.93	0.38
Motivation to Read	3.85*	1.03	4.91*	0.80	4.50*	0.84
Written Sample I No. of Words	2.63*	2.80	3.37*	1.84	1.74	2.95
Written Sample I Clause Index	1.39	0.12	0.49	0.14	0.78	0.18
Written Sample I T-unit length	0.14	0.79	-0.28	1.14	-1.48	0.95
Written Sample II No. of Words	-0.48	4.59	0.91	4.08	-1.81	4.65
Written Sample II Clause Index	-3.27*	0.17	-0.33	0.17	-0.83	0.18
Written Sample II T-unit length	-0.53	1.80	0.13	1.47	-1.44	1.48
Lorge-Thorndike Intelligence Test	0.32	2.41	-0.01	1.92	0.86	2.54
Socio-Economic	0.34	0.28	-2.40**	0.22	-6.36*	0.23
Linguistic Reading Test	Insufficient Observations for Calculations					
Total Reading Time	2.86*	10.16	2.62*	7.51	-0.07	10.09

Minus sign on t-value indicates that the second program has a higher mean
*.01 level of significance
**.02 level of significance
***.05 level of significance

Table 13 T-Test between the *Control* and *EDL-Basal* groups for selected dependent variables blocking on levels of Murphy-Durrell Total Test readiness

	Level I Total Score: 98-118		Level II Total Score: 74-97		Level III Total Score: 16-73	
Dependent Variable	t-value	S.D.	t-value	S.D.	t-value	S.D.
Sex	-0.49	0.06	0.72	0.07	0.44	0.09
Gates-MacGinitie I Vocabulary	3.54*	1.03	4.66*	1.15	2.73*	1.64
Gates-MacGinitie I Comprehension	1.86	0.85	5.07*	0.73	2.11***	2.08
Gates-MacGinitie II Vocabulary	-0.03	0.70	3.06*	1.25	0.23	2.28
Gates-MacGinitie II Comprehension	0.13	0.78	3.44*	1.03	2.15***	1.20
Attitude toward Reading	1.53	0.24	1.63	0.28	2.21***	0.35
Motivation to Read	-0.46	0.67	0.91	0.74	1.22	0.87
Written Sample I No. of Words	1.58	1.84	2.05***	1.67	0.90	2.77
Written Sample I Clause Index	1.78	0.07	3.35*	0.12	1.53	0.18
Written Sample I T-unit length	-0.31	0.56	2.03***	0.96	0.11	0.99
Written Sample II No. of Words	-2.51**	4.48	-1.77	3.67	0.68	3.65
Written Sample II Clause Index	-0.68	0.10	-0.89	0.16	1.02	0.19
Written Sample II T-unit length	0.53	1.19	0.93	1.24	0.73	1.34
Lorge-Thorndike Intelligence Test	2.77*	1.48	2.83*	1.64	1.86	2.31
Socio-Economic	-1.89	0.16	-5.40*	0.16	-4.96*	0.21
Linguistic Reading Test	-1.76	4.57	1.17	4.76	0.58	4.02
Total Reading Time	1.63	5.80	0.16	6.41	0.11	9.14

Minus sign on t-value indicates that the second program has a higher mean
*.01 level of significance
**.02 level of significance
***.05 level of significance

Table 14 T-Test between *McGraw-Hill Programmed Reading* and *SRA Basic Readers* groups for selected dependent variables blocking on levels of Murphy-Durrell Total Test readiness

	Level I Total Score: 98-118		Level II Total Score: 74-97		Level III Total Score: 16-73	
Dependent Variable	t-value	S.D.	t-value	S.D.	t-value	S.D.
Sex	0.27	0.17	0.52	0.15	0.06	0.11
Gates-MacGinitie I Vocabulary	0.64	2.34	-1.94	3.42	1.57	1.82
Gates-MacGinitie I Comprehension	0.32	2.69	-1.20	3.33	-0.07	0.92
Gates-MacGinitie II Vocabulary	0.99	1.69	-0.28	2.69	3.03*	2.33
Gates-MacGinitie II Comprehension	0.51	1.87	1.01	2.70	2.37**	1.70
Attitude toward Reading	2.14***	0.63	0.67	0.51	0.98	0.42
Motivation to Read	2.37***	1.36	1.68	1.16	2.87*	0.79
Written Sample I No. of Words	0.59	5.63	-0.73	4.64	0.54	2.90
Written Sample I Clause Index	1.08	0.47	0.39	0.29	1.14	0.18
Written Sample I T-unit length	-1.83	1.85	-0.31	0.96	1.38	0.78
Written Sample II No. of Words	0.99	8.91	2.50**	7.59	2.45**	8.94
Written Sample II Clause Index	0.38	0.15	2.09***	0.48	1.80	0.20
Written Sample II T-unit length	0.39	1.02	2.10***	3.03	2.43***	1.41
Lorge-Thorndike Intelligence Test	-1.85	4.30	-1.38	3.26	0.36	2.47
Socio-Economic	-1.90	0.39	-0.78	0.29	-0.73	0.17
Linguistic Reading Test	1.97	3.63	-1.90	5.32	2.94*	4.59
Total Reading Time	2.44**	18.61	1.28	14.06	2.42**	9.55

Minus sign on t-value indicates that the second program has a higher mean
*.01 level of significance
**.02 level of significance
***.05 level of significance

Table 15 T-Tests between *McGraw-Hill Programmed Reading* and *Harper & Row Lingusitic Readers* groups for selected dependent variables blocking on levels of Murphy-Durrell Total Test readiness

	Level I Total Score: 98-118		Level II Total Score: 74-97		Level III Total Score: 16-73	
Dependent Variable	t-value	S.D.	t-value	S.D.	t-value	S.D.
Sex	-1.38	0.21	0.52	0.14	-0.87	0.19
Gates-MacGinitie I Vocabulary	0.51	3.33	-1.29	2.55	-1.58	3.27
Gates-MacGinitie I Comprehension	0.28	3.85	-1.72	2.51	-1.55	1.71
Gates-MacGinitie II Vocabulary	-0.29	2.00	-0.72	2.16	-2.53**	4.11
Gates-MacGinitie II Comprehension	1.13	2.36	-0.84	2.20	-0.37	3.03
Attitude toward Reading	3.82*	0.66	2.61**	0.41	-0.86	0.66
Motivation to Read	Insufficient Observations for Calculations					
Written Sample I No. of Words	1.14	6.27	-0.38	3.81	-0.95	4.65
Written Sample I Clause Index	0.49	0.54	0.06	0.25	-0.72	0.31
Written Sample I T-unit length	0.40	1.64	-0.64	0.84	-1.16	1.20
Written Sample II No. of Words	Insufficient Observations		0.33	7.60	1.06	13.96
Written Sample II Clause Index	Insufficient Observations		-1.29	0.48	-1.80	0.26
Written Sample II T-unit length	Insufficient Observations		-0.98	3.01	0.44	1.95
Lorge-Thorndike Intelligence Test	-0.81	5.03	-1.17	3.15	-0.88	4.39
Socio-Economic	1.12	0.17	2.61**	0.48	6.50*	0.27
Linguistic Reading Test	0.17	3.54	-3.24*	3.94	1.34	7.09
Total Reading Time	0.37	22.03	-0.37	12.65	-0.76	16.70

Minus sign on t-value indicates that the second program has a higher mean
*.01 level of significance
**.02 level of significance

Table 16 T-Tests between *McGraw-Hill Programmed Reading* and *Listen, Look, Learn* groups for selected dependent variables blocking on levels of Murphy-Durrell Total Test readiness

	Level I Total Score: 98-118		Level II Total Score: 74-97		Level III Total Score: 16-73	
Dependent Variable	t-value	S.D.	t-value	S.D.	t-value	S.D.
Sex	0.00	0.11	0.57	0.08	1.10	0.11
Gates-MacGinitie I Vocabulary	2.74*	1.64	2.95*	1.65	0.60	1.92
Gates-MacGinitie I Comprehension	1.81	1.72	1.10	1.53	-2.34***	1.00
Gates-MacGinitie II Vocabulary	0.43	0.98	2.58**	1.48	1.40	2.23
Gates-MacGinitie II Comprehension	0.96	1.12	2.06***	1.39	0.90	1.58
Attitude toward Reading	0.29	0.40	0.63	0.29	-0.51	0.42
Motivation to Read	3.43*	1.04	3.72*	0.78	2.94*	0.96
Written Sample I No. of Words	1.37	3.19	2.83*	2.34	1.23	2.74
Written Sample I Clause Index	1.86	0.26	0.56	0.16	0.72	0.18
Written Sample I T-unit length	0.20	0.86	-2.42**	0.62	-1.74	0.82
Written Sample II No. of Words	0.24	5.96	2.58**	3.94	0.61	9.25
Written Sample II Clause Index	-3.25*	0.19	0.48	0.25	-0.53	0.14
Written Sample II T-unit length	-2.80*	1.05	0.10	1.58	-0.89	1.48
Lorge-Thorndike Intelligence Test	-0.77	2.67	-1.43	1.92	-1.15	2.38
Socio-Economic	-0.30	0.31	0.05	0.22	-3.43*	0.20
Linguistic Reading Test	Insufficient Observations for Calculations					
Total Reading Time	-0.23	10.88	-2.50**	7.81	-4.37*	9.28

Minus sign on t-value indicates that the second program has a higher mean
*.01 level of significance
**.02 level of significance
***.05 level of significance

Table 17 T-Tests between *McGraw-Hill Programmed Reading* and *EDL-Basal* groups for selected dependent variables blocking on levels of Murphy-Durrell Total Test readiness

	Level I Total Score: 98-118		Level II Total Score: 74-97		Level III Total Score: 16-73	
Dependent Variable	t-value	S.D.	t-value	S.D.	t-value	S.D.
Sex	-1.29	0.07	0.80	0.08	-0.57	0.10
Gates-MacGinitie I Vocabulary	6.57*	1.25	6.46*	1.35	3.22*	1.51
Gates-MacGinitie I Comprehension	5.89*	1.09	4.61*	1.29	2.51**	0.81
Gates-MacGinitie II Vocabulary	2.06***	0.74	5.71*	1.23	1.24	2.20
Gates-MacGinitie II Comprehension	2.80*	0.84	5.78*	1.21	2.59**	1.46
Attitude toward Reading	1.66	0.29	2.63*	0.27	2.38**	0.39
Motivation to Read	-0.99	0.73	-0.44	0.75	0.09	0.96
Written Sample I No. of Words	-0.03	2.28	1.83	2.11	0.27	2.58
Written Sample I Clause Index	3.05*	0.15	3.23*	0.13	1.47	0.17
Written Sample I T-unit length	-0.16	0.69	1.49	0.54	0.09	0.91
Written Sample II No. of Words	-1.30	5.90	0.00	3.64	2.17***	7.62
Written Sample II Clause Index	-1.01	0.11	0.13	0.23	1.66	0.17
Written Sample II T-unit length	-1.39	0.97	0.85	1.33	1.33	1.35
Lorge-Thorndike Intelligence Test	0.72	1.78	1.95	1.66	-0.28	2.19
Socio-Economic	-2.58**	0.19	-2.19***	0.15	-1.34	0.18
Linguistic Reading Test	0.26	3.22	1.21	5.00	2.50***	6.32
Total Reading Time	-3.38*	6.55	-5.72*	6.68	-4.59*	8.46

Minus sign on t-value indicates that the second program has a higher mean
*.01 level of significance
**.02 level of significance
***.05 level of significance

Table 18 T-Tests between *SRA Basic Readers* and *Harper & Row Linguistic Readers* groups for selected dependent variables blocking on levels of Murphy-Durrell Total Test readiness

	Level I Total Score: 98-118		Level II Total Score: 74-97		Level III Total Score: 16-73	
Dependent Variable	t-value	S.D.	t-value	S.D.	t-value	S.D.
Sex	-1.25	0.27	-0.03	0.20	-0.82	0.21
Gates-MacGinitie I Vocabulary	0.04	4.97	0.89	3.76	-2.49**	3.23
Gates-MacGinitie I Comprehension	0.05	4.80	-0.13	2.42	-1.75	1.47
Gates-MacGinitie II Vocabulary	-0.93	2.42	-0.26	3.06	-4.17*	4.19
Gates-MacGinitie II Comprehension	0.59	2.92	-1.74	2.63	-1.62	3.20
Attitude toward Reading	0.97	1.20	0.88	0.83	-1.16	0.84
Motivation to Read	Insufficient Observations for Calculations					
Written Sample I No. of Words	0.60	6.28	0.36	5.37	-1.14	5.23
Written Sample I Clause Index	-1.62	0.15	-0.70	0.14	-1.50	0.29
Written Sample I T-unit length	0.90	4.48	-0.20	1.20	-1.52	1.63
Written Sample II No. of Words	Insufficient Observations		-13.11*	1.26	-1.41	5.00
Written Sample II Clause Index	Insufficient Observations		3.15**	0.51	1.68	0.49
Written Sample II T-unit length	Insufficient Observations		-4.93*	1.89	-2.23***	1.92
Lorge-Thorndike Intelligence Test	0.70	5.56	0.20	4.11	-0.88	5.39
Socio-Economic	0.00	0.00	0.00	0.00	0.00	0.00
Linguistic Reading Test	-1.53	4.28	-0.51	5.22	-0.95	4.25
Total Reading Time	-2.00	18.52	-2.05	11.08	-2.16***	16.52

Minus sign on t-value indicates that the second program has a higher mean
*.01 level of significance
**.02 level of significance
***.05 level of significance

Table 19 T-Tests between *SRA Basic Readers* and *Listen, Look, Learn* groups for selected dependent variables blocking on levels of Murphy-Durrell Total Test readiness

	Level I Total Score: 98-118		Level II Total Score: 74-97		Level III Total Score: 16-73	
Dependent Variable	t-value	S.D.	t-value	S.D.	t-value	S.D.
Sex	-0.24	0.19	-0.17	0.16	0.80	0.14
Gates-MacGinitie I Vocabulary	0.90	3.35	3.47*	3.31	-0.73	2.35
Gates-MacGinitie I Comprehension	0.79	2.88	3.05*	1.87	-1.99	1.14
Gates-MacGinitie II Vocabulary	-0.73	1.71	1.40	3.27	-1.37	2.87
Gates-MacGinitie II Comprehension	0.07	1.86	0.06	2.37	-1.34	1.95
Attitude toward Reading	-1.41	0.87	-0.23	0.69	-1.03	0.61
Motivation to Read	0.56	0.60	1.27	0.76	1.59	0.35
Written Sample I No. of Words	0.24	4.34	2.79*	3.59	0.54	3.34
Written Sample I Clause Index	-0.15	0.13	-0.15	0.18	-0.45	0.18
Written Sample I T-unit length	1.48	2.40	-0.83	1.43	-1.91	1.31
Written Sample II No. of Words	-1.23	6.04	-1.74	5.09	-2.86*	5.66
Written Sample II Clause Index	-1.52	0.44	-2.46***	0.35	-1.60	0.27
Written Sample II T-unit length	-1.47	2.28	-2.65**	2.34	-2.64**	1.79
Lorge-Thorndike Intelligence Test	1.27	4.66	0.55	3.18	-1.12	3.23
Socio-Economic	0.97	0.67	0.46	0.52	-2.31***	0.25
Linguistic Reading Test	Insufficient Observations for Calculations					
Total Reading Time	-3.13*	15.26	-3.14*	11.92	-5.39*	11.82

Minus sign on t-value indicates that the second program has a higher mean
*.01 level of significance
**.02 level of significance
***.05 level of significance

Table 20 T-Tests between *SRA Basic Readers* and *EDL-Basal* groups for selected dependent variables blocking on levels of Murphy-Durrell Total Test readiness

	Level I Total Score: 98-118		Level II Total Score: 74-97		Level III Total Score: 16-73	
Dependent Variable	t-value	S.D.	t-value	S.D.	t-value	S.D.
Sex	-0.82	0.17	-0.10	0.15	-0.47	0.13
Gates-MacGinitie I Vocabulary	2.19***	3.09	5.78*	2.65	1.19	1.69
Gates-MacGinitie I Comprehension	2.43*	2.27	5.33*	1.86	2.34***	0.90
Gates-MacGinitie II Vocabulary	-0.08	1.72	2.86*	2.72	-1.53	2.85
Gates-MacGinitie II Comprehension	0.74	1.92	1.80	2.37	-0.16	1.59
Attitude toward Reading	-1.20	0.72	0.52	0.69	0.90	0.56
Motivation to Read	-3.15*	1.25	-2.23***	1.02	-3.02*	0.72
Written Sample I No. of Words	-0.68	5.03	1.80	4.02	-0.26	3.29
Written Sample I Clause Index	-0.34	0.16	1.91	0.16	0.22	0.20
Written Sample I T-unit length	1.66	1.97	0.88	1.23	-0.62	1.60
Written Sample II No. of Words	-1.30	12.68	-2.96*	6.42	-1.37	3.91
Written Sample II Clause Index	-0.66	0.25	-2.27***	0.43	-0.26	0.30
Written Sample II T-unit length	-0.76	2.31	-2.94*	1.78	-0.94	1.72
Lorge-Thorndike Intelligence Test	2.28***	4.05	2.07***	3.09	-0.48	3.13
Socio-Economic	0.57	0.45	-0.42	0.25	-0.54	0.21
Linguistic Reading Test	-1.61	3.92	2.46***	6.57	0.54	4.21
Total Reading Time	-5.51*	12.23	-5.41*	10.37	-5.59*	11.08

Minus sign on t-value indicates that the second program has a higher mean
*.01 level of significance
**.02 level of significance
***.05 level of significance

Table 21 T-Tests between *Harper & Row Linguistic Readers* and *Listen, Look, Learn* groups for selected dependent variables blocking on levels of Murphy-Durrell Total Test readiness

	Level I Total Score: 98-118		Level II Total Score: 74-97		Level III Total Score: 16-73	
Dependent Variable	t-value	S.D.	t-value	S.D.	t-value	S.D.
Sex	1.29	0.22	-0.15	0.15	1.35	0.21
Gates-MacGinitie I Vocabulary	0.60	4.67	3.42*	2.38	1.69	3.74
Gates-MacGinitie I Comprehension	0.54	3.80	4.50*	1.33	0.16	1.97
Gates-MacGinitie II Vocabulary	0.55	1.82	2.16***	2.49	3.18*	4.25
Gates-MacGinitie II Comprehension	-0.68	2.34	2.62**	1.80	0.90	2.83
Attitude toward Reading	-2.69**	0.89	-1.67	0.53	0.46	0.76
Motivation to Read	Insufficient Observations for Calculations					
Written Sample I No. of Words	-0.67	4.10	3.17*	2.55	2.47**	3.15
Written Sample I Clause Index	1.33	0.16	0.46	0.16	1.98	0.18
Written Sample I T-unit length	-0.31	1.58	-0.76	1.26	-0.02	1.81
Written Sample II No. of Words	Insufficient Observations		1.50	5.11	-1.14	8.03
Written Sample II Clause Index	Insufficient Observations		1.93	0.38	1.22	0.32
Written Sample II T-unit length	Insufficient Observations		1.36	2.28	-0.20	2.28
Lorge-Thorndike Intelligence Test	0.40	5.08	0.30	3.09	0.23	5.01
Socio-Economic	-1.62	0.83	-3.67*	0.48	-5.34*	0.48
Linguistic Reading Test	Insufficient Observations for Calculations					
Total Reading Time	-0.67	15.91	-1.56	9.51	-1.68	16.66

Minus sign on t-value indicates that the second program has a higher mean
*.01 level of significance
**.02 level of significance
***.05 level of significance

Table 22 T-Tests between *Harper & Row Linguistic Readers* and *EDL-Basal* groups for selected dependent variables blocking on levels of Murphy-Durrell Total Test readiness

	Level I Total Score: 98-118		Level II Total Score: 74-97		Level III Total Score: 16-73	
Dependent Variable	t-value	S.D.	t-value	S.D.	t-value	S.D.
Sex	0.90	0.21	-0.07	0.14	0.56	0.21
Gates-MacGinitie I Vocabulary	1.44	4.53	6.28*	1.91	4.06*	2.47
Gates-MacGinitie I Comprehension	1.64	3.24	7.54*	1.36	3.04*	1.54
Gates-MacGinitie II Vocabulary	1.01	2.08	4.13*	2.08	3.08*	4.26
Gates-MacGinitie II Comprehension	-0.13	2.43	4.75*	1.86	2.70**	1.82
Attitude toward Reading	-2.57**	0.79	-0.69	0.55	2.07***	0.72
Motivation to Read	Insufficient Observations for Calculations					
Written Sample I No. of Words	-1.29	5.60	1.71	3.11	1.42	3.61
Written Sample I Clause Index	1.00	0.18	2.90*	0.14	1.81	0.26
Written Sample I T-unit length	-0.40	1.93	1.22	1.08	0.58	2.55
Written Sample II No. of Words	Insufficient Observations		-0.39	6.43	0.33	5.08
Written Sample II Reading Time	Insufficient Observations		1.48	0.44	1.82	0.41
Written Sample II T-unit length	Insufficient Observations		2.36***	1.73	1.15	2.31
Lorge-Thorndike Intelligence Test	1.12	4.77	1.86	2.00	0.64	5.09
Socio-Economic	-3.16*	0.55	-8.82*	0.24	-5.14*	0.41
Linguistic Reading Test	0.08	2.96	4.14*	4.54	1.06	5.95
Total Reading Time	-2.15***	14.07	-4.00*	8.36	-1.61	16.23

Minus sign on t-value indicates that the second program has a higher mean
*.01 level of significance
**.02 level of significance
***.05 level of significance

Table 23 T-Tests between *Listen, Look, Learn* and *EDL-Basal* groups for selected dependent variables blocking on levels of Murphy-Durrell Total Test readiness

	Level I Total Score: 98-118		Level II Total Score: 74-97		Level III Total Score: 16-73	
Dependent Variable	t-value	S.D.	t-value	S.D.	t-value	S.D.
Sex	-0.92	0.11	0.13	0.10	-1.35	0.13
Gates-MacGinitie I Vocabulary	1.87	2.00	2.63*	1.46	1.97	1.89
Gates-MacGinitie I Comprehension	2.24***	1.46	4.58*	0.93	4.07*	1.07
Gates-MacGinitie II Vocabulary	1.12	0.99	1.93	1.65	-0.15	2.78
Gates-MacGinitie II Comprehension	1.15	1.13	3.13*	1.32	1.65	1.43
Attitude toward Reading	0.87	0.43	1.34	0.39	2.10***	0.54
Motivation to Read	-4.47*	0.96	-1.51*	0.72	-3.27*	0.84
Written Sample I No. of Words	-1.56	2.85	-1.38	1.99	-1.05	2.53
Written Sample I Clause Index	-0.37	0.09	3.14*	0.11	-1.05	2.53
Written Sample I T-unit length	-0.28	0.99	2.73*	0.83	0.90	1.68
Written Sample II No. of Words	-1.09	8.30	-2.75*	3.70	2.13***	5.08
Written Sample II Clause Index	2.23***	0.22	-0.36	0.24	1.55	0.23
Written Sample II T-unit length	0.95	1.68	0.80	1.21	1.70	1.83
Lorge-Thorndike Intelligence Test	1.34	2.49	2.37**	1.97	0.71	2.97
Socio-Economic	-1.26	0.32	-1.30	0.27	1.55	0.29
Linguistic Reading Test	Insufficient Observations for Calculations					
Total Reading Time	-2.65*	7.41	-2.86*	6.53	0.16	10.89

Minus sign on t-value indicates that the second program has a higher mean
*.01 level of significance
**.02 level of significance
***.05 level of significance

cantly higher than the EDL-Basal program. However, significant differences also favored Harper-Row over SRA on the vocabulary subtest and LLL over McGraw-Hill on the comprehension subtest. Data for the Form II subtests significantly favored Harper-Row over all other programs on the vocabulary subtest and over EDL-Basal in comprehension. The Harper-Row results were obtained with an extremely low *n*, however, and must be considered within this framework. The McGraw-Hill and Control groups were also significantly higher than SRA on the vocabulary subtest and higher than SRA and EDL-Basal in comprehension. Significant Linguistic test results favored McGraw-Hill for this level.

The results of the analysis of program comparisons of written composition blocking on Murphy-Durrell readiness levels are summarized in Table 24. Significant interprogram differences were indicated for most measures but no significant program-by-level interactions were found. While no one clearly superior program was evident for any measure or level of writing sample I, those significant results which were found, predominantly opposed the multi-sensory programs of LLL or EDL-Basal. The results of the post-test writing sample II, however, presented an altered pattern. For the highest readiness level, the only significant differences favored EDL-Basal over the Controls in the number of words written and LLL over EDL-Basal for the Clause Index measure. However, in both the middle and low readiness levels, all other programs generally out-performed SRA on the three measures. The McGraw-Hill group was also significantly higher than the Controls in the number of words written.

Examination of the data for IQ, Sex, and socio-economic status by readiness levels revealed few factors outside the programs themselves which might contribute to high or low program performance. Although interprogram variations in these measures, caused by clustering, did occur from level to level, they were not consistent or pertinent to variations in program performance. However, the low level SRA group was found to be the lowest of the groups in unadjusted mean total readiness score and their relatively poorer performance at this level undoubtedly reflects the high relationship of readiness to achievement.

In summary, it was found that when subjects were blocked for pretest level of readiness, McGraw-Hill *Programmed Reading* was favored most often for all reading achievement measures on all readiness levels. In the high and middle levels, it was surpassed significantly by no other program. Results favoring SRA and Harper-Row, which occurred predominantly in the middle and low readiness levels, must be viewed with less confidence because of the low *ns* of these programs, even though only

Table 24 F-ratios for writing sample measures blocking on levels of Murphy-Durrell Total Test readiness

	Writing Sample I			Writing Sample II		
	Number of words	Clause Index	T-unit length	Number of words	Clause Index	T-unit length
Reading Programs (RP)	4.53*	4.10*	1.15	5.73*	4.05**	3.51*
Levels (L)	61.70*	9.68*	12.20*	13.80*	8.26*	11.40*
RP x L	0.86	1.45	1.41	1.19	1.22	0.45

*.001 level of significance
**.01 level of significance

Table 25 F-ratios for reading achievement measures blocking on levels of pre-test intelligence

	Gates-MacGinitie I (Mid-test)		Gates-MacGinitie II (Post-test)		Linguistic Reading Test
	Vocabulary	Comprehension	Vocabulary	Comprehension	
Reading Programs (RP)	4.65*	4.35*	6.23*	4.63*	3.38**
Levels (L)	87.50*	47.50*	69.10*	71.60*	29.20*
RP x L	1.41	1.24	1.05	2.05***	0.55

*.001 level of significance
**.01 level of significance
*** .03 level of significance

the low level SRA group was surpassed on any measure by McGraw-Hill. The Controls, LLL, and EDL-Basal programs were fairly consistently the low groups for the significant differences which were found. The Controls, using traditional basal materials and methods, surpassed only the EDL-Basal group which paired basal materials with periods of mechanized drill. LLL, which pairs the multi-sensory mechanized approach with a "modified linguistic" program, also only surpassed EDL-Basal, its basal counterpart. The LLL and EDL-Basal programs also fared poorly in the analysis of the mid-year writing sample measures. However, these programs were favored in some of the significant differences found in the analysis of post-test writing samples and, in the middle and low readiness levels, it was SRA which was generally out-performed on all three measures.

Effectiveness blocking on pre-test IQ. The results of the analysis of interprogram reading achievement blocking on pre-test IQ levels are summarized in Table 25. Significant reading program and levels effects were found for all the achievement measures but a significant program-by-level interaction occurred for only the Form II Gates-MacGinitie comprehension subtest. Although for each variable fifteen interprogram comparisons were expected for each level, Harper-Row had an insufficient number of subjects in the low IQ level for inclusion in that analysis. Again, only a review of the significant differences is included in this section and complete data may be found in a separate publication of project statistics available through ECRI.

For the high IQ subjects, significant differences favored McGraw-Hill over both the Controls and EDL-Basal on the Form I Gates-MacGinitie subtests and over the Controls, LLL, and EDL-Basal on the Form II subtests. Significant differences at this level also favored SRA over the Controls for the Form I vocabulary subtest, Harper-Row over EDL-Basal for the Form I comprehension subtest, and EDL-Basal over LLL for the Form II vocabulary subtest. No significant differences were found within this level for the *Linguistic Reading Test.*

In the middle IQ level, McGraw-Hill and Harper-Row were significantly higher than both the Controls and EDL-Basal for the Form I test and higher than the Controls and SRA for the Form II subtests. Harper-Row was also favored over LLL and EDL-Basal on the Form II vocabulary subtest. The results of the Linguistic test within this level also favored McGraw-Hill and Harper-Row over the Controls and SRA.

Within the lowest IQ level, no one program was favored consistently across all measures. Harper-Row was not included in this analysis. For the

Form I test results, McGraw-Hill was significantly higher than the Controls and EDL-Basal for the vocabulary subtest and, for the comprehension subtest, the Controls were favored over EDL-Basal. All significant differences of the Form II subtests opposed SRA. McGraw-Hill and EDL-Basal were favored for vocabulary and the Controls, while LLL and EDL-Basal were favored in comprehension. No significant differences were found within this level for the *Linguistic Reading Test.*

The results of interprogram comparisons of written composition blocking on pre-test IQ are summarized in Table 26. Significant program and level effects were indicated for many of the measures but no significant program-by-level interactions were found. For the mid-test writing sample (I), significant differences occurred predominantly in the number of words written but no positive pattern favoring a particular program resulted for the high and middle IQ levels. Instead, a number of programs were favored over LLL in the number of words measure in both levels and over EDL-Basal for the Clause Index measure within the middle level. Within the lowest IQ level, however, LLL was significantly higher than SRA for the Clause Index measure and favored over the Controls, SRA, McGraw-Hill, and EDL-Basal for the measure of average T-unit length. The post-test writing samples (II) again resulted in LLL being predominantly the low group in the number of words written within the highest IQ level. In the middle IQ level, however, SRA was most frequently the low group in the number of words written and was consistently low for all measures within the lowest IQ level.

In examining the data for possible clustering effects which may have influenced interprogram results for a particular level, unique readiness factors were apparent for the middle and low IQ level SRA groups. Their unadjusted means for the Murphy-Durrell Total Test and Letter Names measures were, by a striking gap, the lowest ranking within these levels. Therefore, for this study, the generally poor performance of the middle and low IQ SRA groups must be assumed to be a product of the strong relationship of these readiness factors to achievement rather than a within-program factor. An inverse relationship was found for the EDL-Basal group, however. While this group consistently produced the highest ranking unadjusted mean for all readiness measures on all levels, significantly lower achievement resulted on all levels and for most measures. Thus, in this instance, added confidence is given the conclusion that within-program factors, or possibly unmeasured factors, are acting to produce the relatively lower EDL-Basal results.

Among the other variables possible of clustering on levels, variations

Table 26 F-ratios for writing sample measures blocking on levels of pre-test intelligence

	Writing Sample I			Writing Sample II		
	Number of words	Clause Index	T-unit length	Number of words	Clause Index	T-unit length
Reading Programs (RP)	3.45**	3.95**	1.38	6.04*	3.96*	3.28**
Levels (L)	9.30*	4.14***	9.05*	0.62	2.75	3.57***
RP x L	0.19	0.52	1.13	1.04	1.03	0.36

*.001 level of significance
**.01 level of significance
***.02 level of significance

in the ratio of boys to girls among programs and levels did exist. There was, however, no discernible relationship to program performance. In socio-economic ranking, the greatest variation occurred with the LLL group, which dropped from third in the high IQ level to last, or lowest in status, in the middle and low IQ levels. Although this would indicate a clustering of low status students within these levels, the significant differences which did favor this program occurred, almost entirely, at these lower levels. This may well evidence the relative unimportance of socio-economic status compared to other variables.

In summary, when blocking on pre-test IQ, McGraw-Hill was favored most frequently within the high and middle IQ levels in significant differences resulting from an analysis of reading achievement measures. Although several significant differences also favored Harper-Row within the middle level, the extremely low *n* of this group severely weakens the confidence which may be placed in these results. The low groups in these comparisons were almost without exception the Controls, EDL-Basal, and, with only slightly less frequency, Listen, Look, Learn. Within the lowest IQ level, McGraw-Hill was again favored over the Controls and EDL-Basal for the Form I vocabulary measure. However, SRA was the low group in all significant differences resulting from the Form II subtests. The analysis of aspects of written composition within IQ levels indicated no one superior program. However, the high and middle IQ level LLL groups and the middle and low IQ SRA groups were the significantly lower groups most frequently. Since examination of the data indicated that the middle and low IQ level SRA groups were by far the lowest ranking group in pre-test readiness, their low performance at these levels must be viewed, at this time, as a product of the high relationship of readiness to achievement rather than as a feature inherent in the program. Conversely, within all levels, EDL-Basal was consistently the highest ranking group in mean readiness. This may be construed as giving added confidence to the conclusion that the low achievement performance of the EDL-Basal groups does indeed reflect an inherent program defect.

CHAPTER V

SUMMARY AND CONCLUSIONS

The reported comparative study of first-grade reading instructional methods was supported by the Granite School District Exemplary Center for Reading Instruction. Extending throughout the 1966-67 school year, it was undertaken in 47 classrooms of nineteen district schools and involved 1295 students.

Summary of Design and Procedures

The design reported in this study was based upon the "quasi-experimental non-equivalent control group design" discussed by Campbell and Stanley (1963) in which existing intact classroom units, involved as experimental or control groups, were given pre-tests and post-tests. Specified factors, apt to threaten internal or external validity of the research, dictated procedures within this broad framework. Validity factors of particular importance to this study concern effects resulting from intrasession history, instrumentation, statistical regression, and biases resulting from selection.

Eighteen of the forty-seven participating first-grade classrooms were control groups, randomly selected for use in establishing a norm for district first grades. Basal reading series, typical of those in general district use, served as instructional materials within these classrooms.

The twenty-nine experimental classrooms were those for which the principal and/or teacher had already elected to provide an innovative or "linguistic" reading program prior to the design of this study. The only exception, the SRA Basic Reading Series, was assigned to all of the first grades of a school which had no classes participating as either control or experimental groups. An acknowledged, and rather severe, imbalance in experimental programs resulted, particularly in the involvement of programs in use in only one classroom. Included were twelve classrooms using the Webster Division, McGraw-Hill *Programmed Reading* materials, two using SRA *Basic Reading Series,* one using Harper and Row *Linguistic Readers,* one using i/t/a materials, five using *Listen, Look, Learn* materials, and eight using EDL machines and materials to supplement a basal reading series. One classroom, excluded from this list but included in the research activities as one of the forty-seven classrooms, was omitted from the resulting data due to inability to maintain program fidelity. In addition, the data from the single i/t/a class was excluded for statistical reasons.

As a factor in the internal validity of the research design, fidelity to experimental programs was assigned primary importance and controlled in several ways. All experimental teachers received workshop instruction in their programs during the summer or early in the school year. Members of the research staff who had familiarized themselves with a program(s) also used classroom observation and fidelity rating scales as training tools for monthly workshop meetings with their teachers. A bibliography of suitable supplementary materials which maintain one-to-one sound-spelling relationships was also provided for the "linguistic" programs in which such control was recommended by the publisher. To account for the amount of reading instructional time, a variable which could be reflected in the measured research results, a computer-scorable time log sheet was developed and was marked by control and experimental teachers for a sampling of days throughout much of the year. Mean instructional time per classroom was computed, and served as a dependent variable for intercorrelation with achievement results.

A group reading-readiness and group IQ test were selected for administration as pre-tests during the second and third week of the school year. Skill-oriented pre-tests were administered in February during a mid-year test series. A standardized reading achievement test, a uniformly assigned written composition, and a reading attitude inventory were administered to the entire population. A subsample of each classroom was given an individual oral reading test to assess fluency and word attack differentials. Observation of reading as a free-time activity was undertaken in randomly selected classrooms. Post-tests, administered in mid-May, included alternate forms of the group and individual reading tests, a second written composition, and administration to a random selection of control and linguistic classrooms of an experimental Linguistic Reading test, presently available in only one form.

To assure conformance to the testing procedures provided with each test instrument, all tests were administered by the research staff following training by the project director. Enough test personnel were used to complete each series in a brief time that would avoid differences in the amount of teaching received by the groups prior to testing, and time of day was equalized among the programs. All scoring was done by members of the scoring staff.

Statistical Summary and Conclusions

As it has become evident that a one-for-all approach to the selection of reading instructional materials can no longer be justified, it has become necessary to seek information concerning the relative effectiveness of reading programs with students of varying characteristics. Thus, the statistical emphasis of this study has been twofold. First, the extent of relationship of readiness and pupil/class factors to reading achievement was examined. Secondly, an investigation was made of possible differences in effectiveness among programs, with pupils grouped homogeneously by two of these factors—assessed readiness and IQ.

Relationships to Achievement. The investigation of the relationship of readiness to reading achievement indicated that the Murphy-Durrell Total Test score was most frequently the best predictor of subsequent reading achievement in first grade. However, it was only slightly better than one of its component parts, the Letter Names subtest. The Bloomfield school of linguists have maintained that a pre-instructional knowledge of letter symbols by name is necessary to prevent their hindrance of the process of learning to read larger units. Whether unfamiliarity with letter names and symbols does indeed establish barriers to reading or merely reflects an essential component of overall preparedness, there are important implications for readiness training, particularly in kindergarten. Since neither the Phonemes subtest—sounds in words—or the Learning Rate subtest—the teaching of whole words—resulted in equally high relationships to achievement, the initial emphasis in readiness training may need to be placed on the accurate recognition of individual symbols.

The Lorge-Thorndike DIQ relationships to achievement were considerably lower than the Murphy-Durrell Total Test or Letter Names scores. However, with relationships ranging from .44 to .68, IQ was found to be the most influential of the factors explored for their relationship to a student's readiness level. Both the sex and the socio-economic status of a student proved to have little or no relationship to either their level of pre-instructional readiness or post-instructional achievement.

There were also generally low correlations between readiness or subsequent reading achievement and the measures of either attitude toward reading or the motivation to read as a free-time activity. Unmeasured factors in the classroom, home, or child himself, and not isolated in this study, perhaps combine to contribute to the interprogram variations which occurred in these measures. The average time devoted to the instruction of reading within a classroom bore little or no relationship to achievement results and varied little among programs. However, difficulties arose in the collection of this data which lessen the confidence to be placed in these results.

Low to moderate relationships to both pre-instructional readiness and post-instructional reading achievement resulted from the measures of unaided written composition. In general, the highest relationships to achievement occurred with the measure of number of words written. Sentence complexity and length are undoubtedly tied to language skills which are relatively unaffected by the growth of first grade reading skills. However, in examining correlations between each of the pre-test writing measures and their post-test counterparts, it appeared that the low and inverse relationships for the complexity and length measures indicate that some change had taken place. Interpretation of this change, which for some programs was progressive and for others regressive, would require information unavailable in this study. To assess these aspects of written language control in future research, it is felt that pre-test and post-test analysis of oral language will need to be included for comparison and, in addition, measurement made of classroom language activities extraneous to the reading program. While the intent of the measurements of this study was to look for resultant differential written language control among programs, we have in reality, only a surface examination that could provide highly interesting future research.

The relationships between the unorthodox *Linguistic Reading Test* and the traditional Gates-MacGinitie reading achievement test were extremely high across all programs without prejudice. Thus, it may be assumed that it is unnecessary to administer a special test for "linguistically" trained groups. While a variation in familiar reading vocabulary does exist between the programs, adjustments to the unfamiliar seem to be made with equal facility by both groups.

Effectiveness by levels of readiness and IQ. In this analysis, a 3 (levels) X 6 (Programs) analysis of variance was performed for each of the twenty-one dependent variables, blocking separately on pre-test Murphy-Durrell Total Test readiness and IQ. Reading program differences and program-by-level interactions were examined to determine whether different programs produced differentiated results in terms of measured pre-test performance.

Blocking on readiness, significant program differences were found for both reading achievement and writing measures. Program-by-level interactions were found for only the reading achievement measures. When subjects were blocked on IQ, significant reading program effects were found for the achievement and writing variables but significant program by level interactions occurred for only the Gates-MacGinitie Form II comprehension subtest.

No single reading program was found to be significantly better than all others in all respects or to be uniquely effective or ineffective for students of any given level of pre-instructional readiness or IQ. Significant differences involving a program tended to occur across all levels and without clear discrimination among variables. However, the innovative "linguistic" programs, which stress regular sound-spelling relationships in vocabulary selection, were, particularly in a programmed format, generally superior to the traditional reading programs and the multi-sensory "modified linguistic" program. McGraw-Hill *Programmed Reading* was favored most frequently across both readiness and IQ levels. The low programs in these differences were quite consistently the Control, EDL-Basal, and Listen, Look, Learn groups. Results favoring SRA and Harper-Row over these same low groups were not as numerous but did appear, particularly in the middle and low levels. Although the extremely low *n* of the SRA and Harper-Row groups lessens the confidence in their results, they were not, with some explicable unconformity by SRA, surpassed significantly by McGraw-Hill and must, in this study, be considered essentially equal to it in performance.

The consistently low performance of the traditional basal programs presents evidence of a need for in-depth revision of these programs. Merely adding to existing programs was not evidenced to be sufficient or even profitable. On all levels, the basal programs, Controls, were only better than similar basal programs, EDL-Basal, in which time was taken for the mechanized drill of materials relatively uncoordinated with the programs. The generally low performance of the multi-sensory, mechanized LLL program, which also surpassed only EDL-Basal, gives rise to the question of the merits of mechanizing instruction at all. However, there would be need to investigate the potential performance of the "modified linguistic" aspect of this program, without the multi-sensory devices, before definitive conclusions could be reached.

The tendency of variables to allow a clustering of either their high or low performing subjects within an established readiness or IQ level provided some additional information about the programs as well as the variables. In point, the exceptions to the general performance of the SRA groups occurred as significantly lower results in the middle and low IQ levels. Examination of the data revealed a clustering of low readiness students at these levels which was manifested in mean readiness lower, by a considerable interval, than the other programs. Therefore, the low performance is assumed to reflect the high relationship of readiness to achievement rather than an inherent program characteristic. However, the EDL-Basal groups consistently had the highest mean readiness on all levels and yet were unable to maintain this advantage.

Somewhat limited evidence was also found which would support the low relationship, in this study, of socio-economic status to achievement. The three programs most frequently favored in the results represented the entire high-middle-low socio-economic range. In addition, extreme clustering of low status students was found in the middle and low IQ levels of the LLL program. It was in these levels, however, that the significant differences favoring LLL occurred. It should be pointed out that the socio-economic range of this study included no hard-core slums or "Appalachias" and, therefore, does not reflect the full range of possible differences. Classrooms of several area Title I schools, as well as relatively high socio-economic classrooms, were included, however, and certainly conditions common to the area, and perhaps to a majority of school districts, are reflected in this study.

Limitations of the Study

Many of the limitations to be recognized in the procedures of this study are attributable to insufficient preplanning time. It has become increasingly apparent that many months of thoughtful preparation are necessary to the initiation of adequate comparative studies. Since this was a pilot study intended as a preliminary vehicle for assessing design and procedures, time limitations are perhaps acceptable, but their implications must be delineated.

First in order of importance to the researcher is the grossness of many of the hypotheses stated for testing in the pilot study. For instance, the questions concerning "achievement," while traditional to comparative field studies, would increase in value if more closely defined. There is a tendency toward "encyclopedism," or general survey, in studies of reading as a result of rationalization that it is an exceedingly complex area composed of processes which are difficult to isolate. However, in decisions concerning the hypotheses to be stated, there is need for the construction of a clear outline of educational objectives which may then be transformed to a behavior which can be measured. Only highly pertinent and clearly defined areas of investigation can be justified in modern research design.

A second area of weakness in this study is related to the statement of clearly specific hypotheses. As this was a hastily initiated pilot study, all innovative programs which were readily available in the district were established as experimental groups. Thus, an important requisite to comparative research design, a contrastive analysis of the unique characteristics of programs being considered for research, was necessarily omitted. A detailed profile of similarities and differences would allow for the selection of contrasting programs possessing characteristics important to the hypotheses to be tested. For instance, if presentation or sequencing of one-to-one sound-spelling relationships were basic to stated hypotheses, "linguistic" reading programs contrastive in this area would be selected. It is from this point that the development of the details of the research design should proceed. Use of experimental programs for which there was no contrastive analysis prior to the experimentation considerably lessened the sensitivity of many of the obtained measures. Procedures and instruments were not always pertinent to isolated aspects of the programs. Suggestion can be made that, quite apart from the uncontrollable variables inherent in even coordinated studies, the haziness of the reported results of the Cooperative First Grade Projects may well have resulted from the lack of precise hypotheses and contrastive analysis of programs during pre-experimental planning. The tested hypotheses were largely relatively gross, and aspects of opposing programs overlapped to cloud the results.

Once the hypotheses have been stated and the programs for comparison selected, the most important procedure which affects the validity of the research is randomization. The use of intact classrooms as the experimental unit rather than random assignment to program by individual is highly defensible. No experimental treatment could be considered auto-instructional and requiring evaluation by individual pupil. Treatment was received by the classroom group in toto. However, an important limitation in this study concerned the number of these units involved in the specific experimental programs. In programs represented by only one classroom, teacher variability is quite apt to be reflected in the measured results as an event outside the experimental situation. A solution to this effect is to increase the number of classroom units per contrasted program.

However, the random assignment of the classroom experimental units to both experimental and control programs is of primary importance to the validity of the study. The use of "self-selected" volunteer experimental teachers, as opposed to control teachers randomly selected from available grade-level personnel, could produce a significant uncontrolled variance in teacher competency and attitude. The sensitivity and generalizability of the results could thus be strongly affected. When possible, assignment to program is most effective when made randomly from the entire teacher population. However, it is quite probable that, due to the amount of cooperation that would be required, selection and assignment would necessarily be confined to volunteer teachers, available for assignment to any one of the treatment programs. The resulting reduction in external validity, or generalizability, is balanced by the degree of internal validity which randomization provides.

Both testing procedure and the selection of instruments also require careful attention as important aspects of design implementation. Measurement limitations within this study were varied. The desirable pre-test/post-test design for each measured variable was not always possible, due in part to the lack of alternate forms of some instruments. In addition, limitations were often imposed by aspects within the selected instruments. A detailed contrastive analysis of the programs, paired with a test item analysis from an operational standpoint, should have been used to influence the choice of the tests used. Failure to perform such analysis perhaps contributed to difficulties such as those encountered with the Gates-MacGinitie "achievement" test. A large number of scores occurred both below the norms and at the extreme top of the A level of this test, recommended for first grade use. In retesting the high group with the next level B test of the series, several extreme downward discrepancies appeared in resulting scores. Doubt concerning the comparability of the two tests, whether due to construction, inappropriateness, or administration, made the choice of which scores to use in statistical intercorrelations impossible. It was decided to use only the A level scores in treatment of the data, acknowledging the flatness in the extreme scores.

In addition, two questions arose about procedures in obtaining the measure of written composition. While the motivational techniques were largely satisfactory, control for the factor of auditory memory was lacking. The possibility that some of the children simply did not remember what they were to write about may have had an extremely limiting effect on their compositions. It is felt that motivational procedures should, in future use, be reinforced by the involvement of a related concrete object or picture which would remain on display throughout the composition period.

Questions of intra-scorer reliability may be raised as a result of the use of a single scorer in evaluating the compositions. Though advantages were obtained in the use of one scorer, the length of time required in scoring almost 1300 compositions may well have produced changes in scorer behavior for which we cannot account. In retrospect, it would seem advantageous in the future to thoroughly train two scorers, using compositions drawn from classrooms not involved in the experiment, until they obtain interscorer reliability. All compositions could then be evaluated by two comparable scorers and reliability increased.

In the administration of the Linguistic Reading Test, the intent was to provide for unprejudiced evaluation of reading skills in all programs. However, administrative problems arose in obtaining scores for this test. Directions provided by the test developers recommend administration of all sections during a single test period. The one and one-half hours required for this test was considered too long a period for first grade children, even though the required response was varied in the seven subsections. The directions and examples allowed were also not considered specific enough for those children who were not able to understand what was required after just one example. Although the developers stated children in their study experienced little difficulty with the tasks required by the test (Anastasiow and Hansen, 1966), this was often not the case in the classrooms of this study, even in those which had been "linguistically" trained and had had previous experience with related tasks.

The *San Diego Inventory of Reading Attitudes*, administered in the study, may also have been prejudiced by what is felt intuitively to be a child's natural inclination to select a positive response when given a yes-no choice. Only four "correct" negative choices were included, a situation which could present a rather narrow range in the obtained results. However, the developers of this instrument report a high reliability in the obtained measures and strong relationship to other indicators of attitude toward reading. In future research, additional attitude-related measures, such as numbers of supplementary books independently read, surveyed home reading activity, or adaptation of the free-time activity observations, measured separately in this study, might all be intercorrelated, as well as combined, to develop an "attitude index" measure.

The free-time motivational observation used in the study was intended as only a pilot effort toward developing a measurement of displayed reading behavior. To be meaningful, it is recognized that a pre-test measure of activity selection would be necessary early in the school year. Also, the present check sheet allowed only for the recording of the selection of reading activity. Since, in this study, writing was considered an important related language area, it might also be well in future use to designate and score selection of writing as well.

An uncontrolled problem arose in the teacher's use of the time log sheets. In the present study, explanation and training in log use was limited to written instructions and one after-school meeting. The principles and mechanics of the logs were explained at this meeting and questions arising at the time were answered. However, examination of the sheets as they were returned, suggested that there may have been a degree of diverse implementation of the principles. It is felt that it would be an advantage to future use of the logs to follow the general explanatory meeting with a half-day observation in each classroom by a staff member. While there, the staff member and the classroom teacher would each mark log sheets for comparison and verification of the teacher's understanding of their use. Multiple visits to a classroom might be required in some instances. The value of this technique was realized during the present study but available staff time was too limited to permit the visitations.

Educational Implications

There has been a recent shift in educational thinking which is reflected in the long-range implications of this study. Educational researchers are no longer looking for the one "best" method for teaching reading but, rather, for properties of different reading programs which are differentially effective with children of varying characteristics. It can be predicted that individualized, prescribed instruction will become increasingly efficient. Careful comparative studies, adapting guidelines of this study, may be expected to contribute substantially to this progress.

More immediate effects on education may also result from the reported study. For instance, emphasis was placed on written control of language as an aspect of early instructional practices. Cited literature discussed the developmental trends toward increased written sentence complexity in the progression toward literacy. As a result of contact with this study, it is hoped that increased attention will be given to early instruction in written language skills. It is also hoped that techniques will be adopted which actively teach use of progressively complex sentences for all levels of communication, oral as well as written.

There are also immediate implications for teacher-training in program fidelity through the workshop-observation-workshop techniques of this study. Lerner (1967) has pointed out that discrepancies are common in the implementation of reading programs in the classroom situation. Undirected, this certainly would have been the case with the experimental groups of this study. Study of sequenced observation sheets for individual teachers indicates a change toward closer fidelity as the workshop-observation-workshop cycle was repeated. A supervised introduction of innovative reading programs could well adopt this training technique in preference to the now common loosely controlled program introduction. Joint supervisor-teacher construction of the observation rating sheets can also be recommended to reinforce initial workshop instruction in program philosophy and techniques.

The stated need for narrowly specific hypotheses and program analysis in reading research also has implication for the continuing diagnostic evaluation of individual progress which is undertaken by the classroom teacher. By narrowly defining her instructional goals as well as developing a profile of the instructional aspects of the materials she is using, she would make her evaluation of a pupil's progress much more effective and useful. Teaching *per se* may also be expected to be more effective as subject matter is closely defined and is viewed in relation to overall educational goals.

REFERENCES

Anastasiow, N. J., and Hansen, D. *Linguistic Reading Test,* Palo Alto Unified Schools, 1966.

Anastasiow, N. J., and Hansen, D. Criteria for linguistic reading programs. *Elementary English,* 1967, *44,* 231.

Anderson, J. E. An evaluation of various indices of linguistic development. *Child Development,* 1937, *8,* 62-68.

Bloomfield, L. Linguistics and reading. *Elementary English,* 1942, *19,* 123-30, 183-86.

Bond, G. L., and Dykstra, R. *Report of Coordinating Center for First-Grade Reading Instruction Programs.* Minneapolis: University of Minnesota, 1967.

Campbell, D. T., and Stanley, J. C. Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of Research on Teaching.* Chicago, Rand McNally, 1963. Pp. 171-245.

Cartwright, G. P. Techniques of analysis of written language. Paper read at American Educational Research Association Annual Meeting, Boston, February, 1966.

Chall, Jeanne. Different approaches to beginning reading. In J. A. Figurel (Ed.), *Reading as an Intellectual Activity.* New York: Scholastic Magazines, 1963. Pp. 250-54.

Cook, D. L. The Hawthorne effect in educational research. *Phi Delta Kappan,* 1962, *44,* 116-22.

Downing, J. A. *The Downing Readers.* London: Initial Teaching Publishing, 1963.

Downing, J. A. *The Initial Teaching Alphabet.* New York: MacMillan, 1964.

Downing, J. A. Current misconceptions about i/t/a. *Elementary English,* 1965, 492-501.

Durrell, D. D., and Murphy, Helen. *Speech to Print.* New York: Harcourt, Brace and World, 1964.

Fries, C. C. *American English Grammar.* New York: Appleton-Century, 1940.

Fries, C. C. *Linguistics and Reading.* Chicago: Holt, Rinehart and Winston, 1963.

Fries, C. C. Linguistic approaches to first grade reading programs. In J. F. Kerfoot (Ed.), *Perspectives in Reading No. 5.* Newark: International Reading Association, 1965. Pp. 45-55.

Fries, C. C., Wilson, Rosemary, and Rudolph, Mildred. *Merrill Linguistic Readers.* Columbus, Ohio: Charles E. Merrill, 1966.

Gates, A. I., and MacGinitie, W. H. *Gates-MacGinitie Reading Tests.* New York: Teachers College Press, 1965.

Gibson, Eleanor. Learning to read. *Science,* 1965, *148,* 1066-72.

Gilmore, J. V. *Gilmore Oral Reading Test.* New York: Harcourt, Brace and World, 1951.

Griffin, W. J. A transformational analysis of the language of kindergarten and elementary school children. Paper read at meeting of American Educational Research Association, Chicago, 1966.

Hahn, H. T. Three approaches to beginning reading instruction – i/t/a, language arts and basic readers. *The Reading Teacher,* 1966, *19,* 590-94.

Harris, A. J., and Server, Blanche. The CRAFT Project: Instructional time in reading research. *Reading Research Quarterly,* 1966, *2,* 27-56.

Hunt, K. W. *Differences in Grammatical Structures Written at Three*

Grade Levels, the Structures to be Analyzed by Transformational Methods. Tallahassee: U.S. Office of Education, Cooperative Research Project No. 1998, 1964.

Hunt, K. W. Recent measures in syntactic development. *Elementary English,* 1966, *44,* 732-39.

Lerner, Janet W. A new focus in reading research: The decision-making process. *Elementary English,* 1967, *44,* 236-42.

Lorge, I., and Thorndike, R. L. *The Lorge-Thorndike Intelligence Tests.* Boston, Houghton-Mifflin, 1957.

Mager, R. F. *Preparing Instructional Objectives.* Palo Alto: Fearon Publishers, 1962.

Mazurkiewicz, A. J. i/t/a and TO reading achievement when methodology is controlled. *The Reading Teacher,* 1966, *19,* 606-10.

Murphy, Helen A., and Durrell, D. D. *Murphy-Durrell Reading Readiness Analysis.* New York: Harcourt, Brace and World, 1965.

Rasmussen, D. E., and Goldberg, Lenina. *S.R.A. Basic Reading Series.* Chicago: Science Research Associates, 1965.

Sheldon, W. D. Basal reading approaches. In J. F. Kerfoot (Ed.), *Perspectives in Reading No. 5.* Newark: International Reading Association, 1965. Pp. 28-44.

Skinner, B. F. Teaching machines. *Scientific American,* 1961, *205,* 91-102.

Smith, H. L., Jr. A review of *Let's Read: A Linguistic Approach. Language,* 1963, *36,* 67-78.

Spache, G. D. *Toward Better Reading.* Champaign: Garrard, 1963.

Stauffer, R. G. The verdict: Speculative controversy. *The Reading Teacher,* 1966, *19,* 563-64.

Stern, Catherine, and Gould, Toni. *Structural Reading Series.* Syracuse: L. W. Singer, 1963.

Stern, Catherine, and Gould, Toni. *Children Discover Reading.* Syracuse: Random House/L. W. Singer, 1965.

Stewart, Rebecca W. i/t/a after two years. *Elementary English,* 1965, *42,* 660-65.

Stratemeyer, Clara G., and Smith, H. L., Jr. *The Linguistic Readers.* Evanston: Harper & Row, 1965.

Sullivan, M. W., and Buchanan, Cynthia. *Programmed Reading.* St. Louis: Webster McGraw-Hill, 1963.

Tanyzer, H. J., and Alpert, H. *Effectiveness of Three Basal Reading Systems on First Grade Reading Achievement.* Hempstead, New York: Hofstra University, 1965.

Tanyzer, H. J., and Alpert, H. Three different basal reading systems and first grade reading achievment. *The Reading Teacher,* 1966, *19,* 636-42.

Tanyzer, H. J., and Mazurkiewicz, A. J. *Early-to-Read i/t/a Program.* New York: Initial Teaching Alphabet Publications, 1966.

Tinker, M. A. Devices to improve speed of reading. *The Reading Teacher,* 1967, *20,* 605-609.

Vogt, W. C., Lane, K. B., Kellogg, R. E., and Norin, W. W. *Inventory of Reading Attitude.* San Diego County Schools, 1965.

APPENDIX A
Samples of Fidelity Rating Scales

S.R.A. Basic Reading Series
Programmed Reading
Look, Listen, Learn
Basal Readers

S.R.A. BASIC READING SERIES

Constructed by Myra H. Castner

Teacher Name: ____________________ Rater: ____________________
Book: ____________________ Date: ____________________
Pages: ____________________ No. of children observed: ____

	(Circle one)	Comments:
A. Classroom speech	0 1 2	
B. Mastery of graphic symbols	0 1 2	
No. of errors ________		
C. Alphabet Book	0 1 2	
D. Discovery method	0 1 2	
E. Linguistic word attack	0 1 2	
F. Pacing of materials presentation	0 1 2	Errors:
No. of errors ________		
G. Instruction sequence observed __, __, __, __, __ (Indicate sequence)		
1. Chalkboard presentation	0 1 2	
2. Word list exercises	0 1 2	
3. Sentences and stories	0 1 2	
4. Silent rereading	0 1 2	
5. Workbook exercises	0 1 2	

SCORING EXPLANATION

An 0 is circled to indicate a behavior not observed or which had no chance to occur. A rating of 2 indicates perfect adherence to the program of the S.R.A. Reading Series. A rating of 1 indicates deviation from any aspect as designated below. The specific deviation(s) is to be noted under the comment column.

A. *Classroom speech*

1. The teacher attempts to alter the speech patterns of herself or of the student(s) during the reading lesson. The presentation of spelling patterns in word lists or sentences is used as an opportunity to correct or change dialect variations or mild speech impediments.

2. In speaking or reading aloud, the teacher's voice reflects her everyday speech. There is no attempt to change a child's varying dialect and/or speech impediment while she is teaching him to read.

B. *Mastery of graphic symbols*
   1. Some children indicate, during the reading period, a lack of skill in names and shapes of several individual letters (number to be noted) and yet have either not used or have moved beyond the Alphabet Book. (Check this status with teacher at end of observation.)
   2. All children who require alphabet training are working in the Alphabet Book (not necessarily to the exclusion of participation in a Series reading book).

C. *Alphabet Book*
   1. Children who need alphabet training are working in the Alphabet Book. However, the pace is kept very slow and the pages repeated for mastery. There is no use of supplementary activities (e.g., chalkboard exercises or alphabet games).
   2. All children who require alphabet training are working in the Alphabet Book (not necessarily to the exclusion of participation in a Series reading book). The pace is rapid enough to maintain interest, most of the children are having largely successful experiences, and varied supplementary activities are utilized.

D. *Discovery method*
   1. Rules, typical of a phonics program, are given by the teacher to "explain" sound-spelling relationships and/or reference is made of word families, such as the "at" family or "it" family, rather than developing concepts through the use of whole words in the word lists.
   2. The teacher presents the word lists and stories in the order designated by the materials and introduces, reviews, and contrasts whole words.

E. *Linguistic word attack*
   1. Isolated sounds of letters are taught and then blended together to form words. For example, "buh" for *b,* "cuh-a-buh" for *cab* and/

or "muh-an" to teach *man.* The teacher also directs the children to look to context, configuration, or picture clues to attack difficult words.

2. The teacher spells out the difficult word (other than exceptional words, which are simply prompted) with vocal emphasis on the error (e.g., "No, this is not *p*-a-t but *b*-a-t.") If necessary, the child is helped additionally with a difficult word in the context of patterns or words he knows. For example:
   "Suppose a child cannot read *rag* . . . Write *ran* on the chalkboard, spell it, and ask the child whether he can read it. If he can, tell him it begins the same way as rag (r-a-g) . . . if he cannot read *ran,* say the word in an ordinary speaking voice and without isolating the /r/ or exaggerating it. If the child still does not succeed, check on the final letter-sound relationships by determining whether the child can read other words on the word list. If he does know other words on the list, he can be told how to read rag, but the teacher should note the difficulty with the introductory r and should watch for a chance to confront the child with new words that begin with r." (1, p. 60)

F. *Pacing of materials presentation*

1. The teacher has taken the children so rapidly through the materials that there are many errors being made in oral reading or the children have remained on each section until all words can be read without error. Interest appears to be low and the words simply memorized as sight words. (Note errors made under comments.)
2. The teacher has all children at their appropriate level. Each child is making "a few mistakes when reading aloud and those occur mostly in words in the pattern being presented in the section he is reading." (1, p. 29) Interest remains high in the group during the entire reading lesson and there is a feeling of success for each child.

G. *Instructional sequence observed*

1. *Chalkboard presentation*
   a. One or more elements are omitted.
   b. Review lists, pattern contrasts, the new word list, sentences, and chalkboard exercises and games (e.g., the missing letter game) are included in the chalkboard presentation.

2. *Word list exercises*
    a. One or more elements are omitted.
    b. After the first word list and the first word in each of the following lists, the children are asked to try to figure out the new words for themselves. They then hear and read the words more than once in order to realize pattern relationships. The teacher makes certain the meaning of each word is understood.

3. *Sentences and Stories*
    a. Content questions and discussion of pictures are a large part of the reading of sentences and stories. Reading with expression is stressed. The teacher provides no motivation toward the reading of the selection and allows complicated word attack procedures (beyond the spelling of the difficult word) to interrupt the thought and interest of the story.
    b. The teacher motivates the reading by building background and adding to the story line when necessary (in the early, limited vocabulary stories). Content questions and discussions of pictures are not a part of the lesson, and word attack techniques do not go beyond the spelling of the word. If this is not sufficient, the child is prompted in order to avoid interference with the flow of the reading. Reading with expression is not stressed but is allowed to evolve naturally with proper teaching and increased reading skill.

4. *Silent rereading*
    a. Silent rereading is omitted as an activity.
    b. Silent rereading is assigned to be done upon return to desks after the group reading period.

5. *Workbook exercises*

PROGRAMMED READING SERIES

Constructed by LaVonne Noren

Teacher Name: ______________
Book: ______________
Pages: ______________
Rater: ______________
Date: ______________
No. of children observed: ______

	(Circle One)	Comments:
A. Presentation of Program		
1. Sound-Symbol	0 1 2	
2. Word Discrimination	0 1 2	
3. Word Formation	0 1 2	
4. Dictation	0 1 2	
B. Tutorial-Self Pacing Method	0 1 2	
C. Social-Emotional Climate of the Program	0 1 2	
D. Mastery of the Material	0 1 2	

RATING SCALE

A. PRESENTATION OF THE PROGRAM

The teacher discusses the use of the programmed reading books in his own organization of the materials, making little or no use of the manual instructions and significantly adds or omits from suggested instructions.

The teacher helps the student to establish correct work habits by presenting to the entire class at one time the text written in the manual concerning the introduction and use of the programmed reading books.

1. Sound-Symbol

   a. The sound-symbol preparation is not given *before* work is begun on a new unit, but rather in a sporadic or unorganized manner. The sound alphabet cards are not used regularly with the lessons.

   b. The sound-symbol preparation is given to the entire class at one time before the fastest students begin working on a new unit. Sounds are referred to as the "m-sound, n-sound," etc. Periodic reviews using the sound-symbol and alphabet cards are given.

2. Word Discrimination

   a. There is no definite attempt to work with the class in pointing out similarities and differences in "like" words.

   b. An exercise in given every day to help the student in word recognition and spelling by pointing out exact differences between such words as pin, pan, pen, etc.

3. Word Formation

   a. Little or no attempt is made to give students exercises in filling in missing parts of words or group words to emphasize phonetic value of the letters.

   b. Spelling skills are developed by giving the students work in filling in missing parts of words. Words have been grouped to emphasize phonetic values of the letters. These exercises occur at least twice a week following exercises in word discrimination.

4. Dictation

   a. Dictation is not in evidence as an entire class exercise. If some dictation is given, the student does not have the opportunity of immediately correcting his work.

   b. Dictation is given to the class at least twice a week. It begins with a few short familiar words and, as the year progresses, dictation is extended to include several short sentences and, finally, paragraphs. "Feed-back" is in evidence by immediately giving the student the correct answers.

B. Tutorial–Self Pacing Method

1. Students work only on a group basis. All members of the group progress in the readers at the same rate. The teacher constantly stops the class to give comments and explanations. There is little evidence of individual tutoring with the students.

2. The students pace themselves individually in working through the readers. No effort is made by the teacher to hold the faster students back or to force the slower ones to cover material which they don't understand. The teacher works on a tutorial basis, observing each student's work and giving special attention whenever test results indicate that it is advisable.

C. Social-Emotional Climate of the Program

1. The teacher often fails to respond with praise to the accomplishments of the students. A general feeling of frustration or dislike for the program is prevalent. Students are admonished regularly not to "cheat" by removing the slider before responding to the question. Other major reading programs are also being used by the teacher with the effect of minimizing the programmed readers.

2. The teacher gives praise often to the students even if many of the answers given are incorrect. Looking at the answer before the response is made by the student is not called cheating but rather is discouraged by explanation. An attitude of relaxation and fun is promoted by the teacher.

D. Mastery of the Material

1. The student is allowed, without a check of mastery by the teacher, to go on to the next book whenever he completes the previous book. If a child is found to be having difficulty, he is given the same material to go over again with no analysis as to the "problem areas."

2. The student has an 80% mastery on the tests before he is allowed to attempt more difficult material. If the child performs poorly on a test, he is given supplementary material and tutored through the first of the next unit.

LOOK, LISTEN, LEARN MULTIMEDIA
Approach to the Teaching of Reading
Constructed by Maurice Wilkinson

Yes No

I Perceptual Accuracy and Visual Efficiency
- A. Tach-X Accuracy Training
  - 1. Timed exposures (.1 - .01 of a second) (Symbol elements, numbers, letters). Immediately after each exposure, the pupils name, draw, or write what they have seen. Use the Look & Write, eyehand coordination workbook. *Single elements.* ____ ____

    Alert
    Flash
    Response
    Check
  - 2. Timed exposures (.1 - .01 of a second) (Series of letters or numbers). ____ ____
    In each exposure recall elements in order in left-to-right fashion. One more element per line than can be seen and retained by the "best" pupil.

    Alert
    Flash
    Response
    Check
  - 3. Never Repeat Exposures ____ ____
  - 4. Check Responses after each exposure ____ ____
  - 5. Daily (5-10 minutes 20 exposures). ____ ____
    Continued until almost all children retain approximately 3-5 numbers or letters accurately.
- B. Accelerated Discrimination Training (Controlled Reader)
  - 1. 5-25 letters per line ____ ____
  - 2. 15-90 lines per minute ____ ____
    (Left-to-right manner).
    Goal is to develop each child's ability to accept projection rates at above 35-40 lines per minute. Increase speed when 75% accuracy reached by three-quarters of the children.
- C. Controlled Reader Motility Training
  - 1. Use the guided slot ____ ____
  - 2. Rates of presentation. ____ ____
    Any rate that approaches 90 lines per minute if the group responds with 75% accuracy in

counting the number of times a given symbol, number or letter occurs.

3. Daily training for 10 minutes. ____ ____
   Continued until all children have mastered projection rates which are above 90 lines per minute.

II Building Experiences

A. Discussions & presentation in preparation for the Aud-X story (either group or class activity) ____ ____

B. Activities directed toward evolving sympathy for, or understanding of the character or situation to be encountered in the Aud-X story. ____ ____

III Skill Building

A. Five kinds of instruction over a four-to-seven day period in the following order:

1. Aud-X introduction—Story Mode ____ ____
   (correlated with Aud-X workbook)
2. Aud-X Word instruction—Word Study Mode (correlated with Aud-X workbook)
3. Tach-X Word Recognition Training ____ ____
4. Controlled Reader Fluency Training-Story ____ ____
   Repeat lessons when necessary because of unsatisfactory performance, absences, or children request. Additional workbooks would be needed.
5. Grouped according to progress and assigned a particular Aud-X center ____ ____

B. Tach-X Word Recognition Training

1. Words flashed at speed B (1/10 second and A 1/100) second). ____ ____
   (Words flashed after child has made a guess what the word would be to complete sentence)
2. Words are used to complete teacher-dictated sentences or workbook sentences from which these words have been omitted. ____ ____

3. Dictate a check at end of lesson and child checks it himself at the end of his workbook. ____ ____
4. Following teacher-presented Tach-X lesson, children complete independent pages in their workbooks. ____ ____
5. Teacher decides from the responses of children to the flashed words whether or not the children need to be recycled through the Aud-X portions of the cycle for reteaching. ____ ____
6. Never be nearer than six foot to the screen. ____ ____

C. Controlled Reader Processing Training

1. Receive a single brief impression of a word. (At no time in this training are children permitted more than a single eye-stop or fixation per word) ____ ____
2. Short High-Speed Bursts—60 to 100 lines. Projected at speeds six times greater per minute than the usual Controlled Reader Rates ____ ____
3. Teacher guides discussions between each "burst." ____ ____
4. Projection and discussion requires approximately 5-8 minutes, maximum. ____ ____
5. Projection distance of approximately 8 feet. ____ ____

D. Controlled Reader Fluency Training

1. Use moving slot. ____ ____
2. No opportunity to regress or reread. ____ ____
3. Children use study guide to preview story. ____ ____
4. Students independently check in the study guide for comprehension after they complete each story. ____ ____
5. Same group of children should use the Aud-X and Controlled Reader. ____ ____
6. Appropriate rate—challenging without confusing, losing their place; maintains attention. ____ ____
7. Comprehension maintained by all at 70% or better. ____ ____
8. Projection distance of approximately 8 feet. ____ ____

9. Minimum goal of 140-160 words per minute for all children. ____ ____
10. Review lessons should be read 10 lines per minute faster than the day's new film is read. ____ ____
11. Scheduled the last day of the cycle. ____ ____

E. Aud-X Training

First of two-day Aud-X Training

1. Story Mode—story introduced, children listen with headsets on (5-8 minutes). ____ ____
2. Workbooks accompany Story Mode. ____ ____

Second of two-day Aud-X Training

3. Word Study Mode (15-20) minutes). ____ ____
4. Workbooks accompany Word Study Mode. ____ ____
5. Aud-X lessons are self-instructional. ____ ____
6. Two Aud-X Centers. ____ ____
7. Six head sets to each Aud-X Center. ____ ____
8. No child farther than 3½ feet from Aud-X screen. ____ ____
9. At the end of the Aud-X training, children complete a skill sheet, word study sheet or story sheet—or all three if appropriate. ____ ____

IV Individualized Reading

A. Literature Program

1. At the end of each of the cycles (excluding the first three cycles) up to the twenty-first a sampler is read by each pupil. ____ ____
2. Questions and study topics accompany each booklet. ____ ____
3. At the end of the twenty-first cycle children begin to read, after every cycle, a poetry unit. ____ ____
4. Study questions for each story or poetry unit (anthologies). ____ ____
5. Carousel Books (classroom library). ____ ____
   Child reads books and evaluates his understanding, interpretation, and appreciation by use skills—check card taken from the file (second half year)

OBSERVATION SCALE

Basal Readers
Constructed by Myra H. Castner

Teacher Name: ______________________ Rater: ______________________
Reading Series: ______________________ Date: ______________________
Book: ______________________ No. of Children observed: ______
Pages: ____

(Circle one) Comments:

I. Grouping 0 1 2

0 - Unobserved or no chance to occur.
1 - Some of the children in the group are reading either far above or below their reading level.
2 - The grouping is such that every child may succeed yet no child is unchallenged.

II. Word Attack method 0 1 2
(Note use of whole word or isolated sounds)

0 - Unobserved or no chance to occur.
1 - Difficult words encountered during reading group activities are merely prompted by the teacher.
2 - Child is aided in decoding difficult words encountered in reading group activity, using word attack methods noted.

III. Independent work 0 1 2

0 - Unobserved or no chance to occur.
1 - Unrelated to building of curriculum-based skills but consists of activities that are merely mechanical.
2 - Involves a meaningful, skill-related activity which will maintain itself with a minimum of teacher direction.

IV. Sequence of reading activities
(Note by letter)

____, ____, ____, ____, ____

A. Introduction of vocabulary
(Note method used)
B. Silent reading
C. Oral reading
D. Work attack skills
(Note skill(s) taught)
E. Work book activity

APPENDIX B

Letter Explaining Use of Time Log and a Copy of a Log Sheet

To: Teachers Participating in the First Grade Reading Project
From: Myra Castner

Attached you will find a sample log. Use a soft No. 2 black pencil. Other colors or ink will not record.

1. Write in your name and school at the top.
2. Fill in your number and school number by blocking in the spaces.
3. Fill in the date by blocking in the space(s) for the day, the month (Ap=April), and the year.
4. Fill in the total number of minutes you direct each activity.
5. Going across each recording area, block in the box segment indicating digits for that time interval. The hundreds digit is always the upper level (when needed), the tens digit is the middle level (or upper when no hundreds digit), and the units digit is the lower level.
6. The numbers listing text material are the numbers you will fill in for each program you are using.

Examples:

   a. If you are using E.D.L. *and* Speech-to-Print, block in the 6 digit after E.D.L. and the 8 digit after Speech-to-Print.
   b. If you are using S.R.A., block in the 3 digit after S.R.A.

Please be sure that any line in the recording area has only *ONE* blocked-in segment. Stray marks and lines will interfere with processing efficiency.

GRANITE SCHOOL DISTRICT
EXEMPLARY CENTER FOR READING INSTRUCTION
SALT LAKE CITY, UTAH
DAILY TEACHER LOG

TEACHER ______________________ SCHOOL ______________________

PLEASE MAKE DENSE BLACK MARKS IN APPROPRIATE SPACES WITH A NO. 2 PENCIL. PLEASE ERASE MARKS THAT ARE NOT SIGNIFICANT.

TEACHER	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
SCHOOL	100	200	300	400	500	600	700	800	900
	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
DAY	10	20	30						
	1	2	3	4	5	6	7	8	9
MONTH	SE	OC	NO	DE	JA	YEAR	67	68	69
	FE	MR	AP	MY	JU		70	71	72

READING ACTIVITIES

GROUP I	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
GROUP II	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
GROUP III	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
GROUP IV	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
GROUP V	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
TOTAL TIME IN GROUP INSTRUCTION	100	200	300						
	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9

TEACHER DIRECTED READING

SIGHT WORD DRILL	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
PHONIC ACTIVITY	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
EXPERIENCE CHART	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
COMPREHENSION	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
READING ALOUD IN TEAMS	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
OTHER	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
INDIVIDUAL READING	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9

TEACHER DIRECTED SUPPORTIVE ACTIVITIES

LISTENING TO STORIES	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
LISTENING TO POETRY	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
CREATIVE WRITING	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
AUDIO-VISUAL ACTIVITY	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
DRAMATIZATION	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
OTHER	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
ARTWORK WITH READING—INDIVIDUAL	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
READING IN SOCIAL STUDIES	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
READING IN SCIENCE	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9

TEXT MATERIALS

AMERICAN BOOK	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
E.D.L.	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
GINN	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
HARPER ROW LINGUISTIC	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
HOUGHTON—MIFFLIN	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
ITA	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
LISTEN LOOK AND LEARN	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
McGRAW-HILL PROGRAMMED	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
MACMILLAN	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
SRA	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9
SPEECH IN PRINT	10	20	30	40	50	60	70	80	90
	1	2	3	4	5	6	7	8	9

IBM H95331

APPENDIX C

Letter of Explanation
and
Copy of Bibliography of Supplementary Reading Materials

December 15, 1966

To: Teachers of Linguistic Reading Programs

Attached is a short bibliography of supplementary books which may be used for your library shelf. I have only included books which are available, in limited number, through the Reading Center Library. Each of you is welcome to check out books from this library although you might first check other sources more convenient for you.

Many of the books listed are the Readers of linguistic programs and may be involved in our research project. They are included in the list because they closely control the presentation of spelling patterns and would provide the best opportunity for a successful independent reading experience for your children. Their use, strictly as independent library material, will not contaminate the instructional program you are using.

A few of the books listed are rather marginal in their control of vocabulary but are included as the best currently available. As more books are found to be suitable, additions to this bibliography will be sent to you.

I hope you find this bibliography helpful in providing successful reading experiences for your class.

Myra Castner
First Grade Reading Research
Project

BIBLIOGRAPHY OF SUPPLEMENTARY LINGUISTIC READING MATERIALS

AUTHOR	TITLE	PUBLISHER
Berg	The Wee Little Man	Follett
Berg	Little Red Hen	Follett
Brothers	Sad Mrs. Sam Sack	Follett
Bulette	The Elf in the Singing Tree	Follett
Fries	Merrill Linguistic Readers Books 1 through 4	Merrill
McCracken	Basic Reading Preprimer	Lippincott
McCracken	Basic Reading Primer	Lippincott
Moore, Susan	Mad Sam (B-1)	SRA
Moore, Susan	Jet, the Pet (B-2)	SRA
Moore, Susan	The Wet Pet (B-3)	SRA
Moore, Susan	Pig Wig and the Ham (B-4)	SRA
Moore, Susan	Tim Gets a Hit (B-5)	SRA
Moore, Susan	A Dog in the Fog (B-6)	SRA
Moore, Susan	In Van's Hut (B-7)	SRA
Moore, Susan	The Bug's Jug (B-8)	SRA
Moore, Susan	Al Tugs a Log (B-9)	SRA
Moore, Susan	The Wig (B-10)	SRA
Moore, Susan	Was Meg Mad? (B-11)	SRA
Moore, Susan	Ben's Pals (B-12)	SRA
Moore, Susan	Ten Tan Bugs (B-13)	SRA
Moore, Susan	A Pig in a Van (B-14)	SRA
Moore, Susan	The Zigzag Hen (B-15)	SRA
Moore, Susan	Dad and the TV (B-16)	SRA
Moore, Susan	Ken's Pet Pup (B-17)	SRA
Moore, Susan	Dig a Bit (B-18)	SRA
Moore, Susan	The Fat Cat (B-19)	SRA
Moore, Susan	Red Dot (B-20)	SRA
Moore, Susan	The Bug (B-1)	SRA
Moore, Susan	Bob's Job (B-22)	SRA
Moore, Susan	Let's Get (B-23)	SRA
Moore, Susan	Sid and Gus (B-24)	SRA
Moore, Susan	Fun in a Tub (B-25)	SRA
Moore, Susan	Fox in a Fix (B-26)	SRA
Moore, Susan	I Can, I Can! (B-27)	SRA

AUTHOR	TITLE	PUBLISHER
Moore, Susan	The Fix-it Man (B-28)	SRA
Moore, Susan	Is a Fox a Pal? (B-29)	SRA
Moore, Susan	The Ox Had No Zip (B-30)	SRA
Moore, Susan	Will's Cab (C-1)	SRA
Moore, Susan	Mix-Up (C-2)	SRA
Moore, Susan	Miss Bell (C-3)	SRA
Moore, Susan	The Bell (C-4)	SRA
Moore, Susan	A Bag of Eggs (C-5)	SRA
Moore, Susan	Max's Luck (C-6)	SRA
Moore, Susan	Dick's Pig (C-7)	SRA
Moore, Susan	Pick it and Pack it (C-8)	SRA
Moore, Susan	A Sick Doll (C-9)	SRA
Moore, Susan	Jack's Nap (C-10)	SRA
Moore, Susan	Lend a Hand (C-11)	SRA
Moore, Susan	Mel and the Fox (C-12)	SRA
Moore, Susan	Is it a Gull? (C-13)	SRA
Moore, Susan	The Hunt (C-14)	SRA
Moore, Susan	The Fun Box (C-15)	SRA
Moore, Susan	Mrs. West's Rest (C-16)	SRA
Moore, Susan	A Cat's Job (C-17)	SRA
Moore, Susan	The Risk (C-18)	SRA
Moore, Susan	Six Hands (C-19)	SRA
Moore, Susan	A Big Wind (C-20)	SRA
Moore, Susan	Pat's Test (C-21)	SRA
Moore, Susan	Wags and the Vet (C-22)	SRA
Moore, Susan	An Elf to Help (C-23)	SRA
Moore, Susan	The Lamp Helps (C-24)	SRA
Moore, Susan	Camp Pants (C-25)	SRA
Moore, Susan	Ask a Clam (C-26)	SRA
Moore, Susan	What Am I? (C-27)	SRA
Moore, Susan	The Club (C-28)	SRA
Moore, Susan	The Elk Himself (C-29)	SRA
Moore, Susan	In the Tent (C-30)	SRA
Moore, Susan	Plums on the Steps (C-31)	SRA
Moore, Susan	Redlegs and the Ducks (C-32)	SRA

AUTHOR	TITLE	PUBLISHER
Moore, Susan	I Would if I Could (C-33)	SRA
Moore, Susan	The Help Club (C-34)	SRA
Moore, Susan	Stub Helps (C-35)	SRA
Moore, Susan	What to Pack for a Trip (C-36)	SRA
Moore, Susan	Bad for the Crab? (C-37)	SRA
Moore, Susan	Pellmell and the Hill (C-38)	SRA
Moore, Susan	Fred's Frog (C-39)	SRA
Moore, Susan	Mrs. Pick's Hat (C-40)	SRA
Moore, Susan	Bess, the Mess (C-41)	SRA
Moore, Susan	A Trip to the Cliff (C-42)	SRA
Moore, Susan	Dad's Truck is Stuck (C-43)	SRA
Moore, Susan	Is a Dress a Mop? (C-44)	SRA
Moore, Susan	Scruff Gets Struck (C-45)	SRA
Moore, Susan	Ken's Gang and the Bobcat (D-1)	SRA
Moore, Susan	Did Bing Sing? (D-2)	SRA
Moore, Susan	Bells Are for Ringing (D-3)	SRA
Moore, Susan	Camping (D-4)	SRA
Moore, Susan	The Sled Sped On (D-5)	SRA
Moore, Susan	What a Trick! (D-6)	SRA
Moore, Susan	Matt, the Bat (D-7)	SRA
Moore, Susan	Big Hank's Pranks (D-8)	SRA
Moore, Susan	The Trick Ring (D-9)	SRA
Moore, Susan	Stan and His Dad, the Junkman (D-10)	SRA
Moore, Susan	Who Has a Wish? (D-11)	SRA
Moore, Susan	Ben Gets Trash Hash (D-12)	SRA
Moore, Susan	A Gift for Pam (D-13)	SRA
Moore, Susan	A Fish Wish (D-14)	SRA
Moore, Susan	Crash! Bang! Smash! (D-15)	SRA
Moore, Susan	Let's Just Think (D-16)	SRA
Moore, Susan	A Trip in a Tub (D-17)	SRA
Moore, Susan	Just a Prank? (D-18)	SRA
Moore, Susan	A Thinking Dog (D-19)	SRA
Moore, Susan	The Elf Who Was Glum (D-20)	SRA
Moore, Susan	Why I Went to Bed (D-21)	SRA
Moore, Susan	Don't Chat So Much! (D-2)	SRA

AUTHOR	TITLE	PUBLISHER
Moore, Susan	Mr. Fix It's Odd Job (D-3)	SRA
Moore, Susan	Not Much for Lunch (D-24)	SRA
Moore, Susan	Guns in the Hills (D-5)	SRA
Moore, Susan	The Red Dutch Chest (D-6)	SRA
Moore, Susan	The Hunting Dog (D-7)	SRA
Moore, Susan	Class Is Not for Ducks (D-28)	SRA
Moore, Susan	Go to Bed (D-29)	SRA
Moore, Susan	The Jumping Match (D-30)	SRA
Moore, Susan	The Inn of the Seven Robbers (E-1)	SRA
Moore, Susan	The Butter Bug (E-2)	SRA
Moore, Susan	Grandmother's Grasshopper (E-3)	SRA
Moore, Susan	The Seventh Buttercup (E-4)	SRA
Moore, Susan	Grandfather Whiskers (E-5)	SRA
Moore, Susan	The Pickle Picnic (E-6)	SRA
Moore, Susan	The Riddle (E-7)	SRA
Moore, Susan	The Dashing and Proper Ant (E-8)	SRA
Moore, Susan	The Poppy Quilt (E-9)	SRA
Moore, Susan	The King of Gobble Penny (E-10)	SRA
Moore, Susan	Tea at the Beach (E-11)	SRA
Moore, Susan	A Bottle of Beans (E-12)	SRA
Moore, Susan	The Blackberry Jelly Elf (E-13)	SRA
Moore, Susan	The Land of Sweet Dreams (E-14)	SRA
Moore, Susan	The Greedy Peacocks (E-15)	SRA
Moore, Susan	Daisy Paints for Fun (E-16)	SRA
Moore, Susan	The Pet Shop Riddle (E-17)	SRA
Moore, Susan	Moon Dreams (E-18)	SRA
Moore, Susan	Clay's Wish (E-19)	SRA
Moore, Susan	The Ugly Nickel (E-20)	SRA
Moore, Susan	The Sick Owl (E-21)	SRA
Moore, Susan	Miss Brown's Bottle (E-22)	SRA
Moore, Susan	Grandfather's Luck (E-23)	SRA
Moore, Susan	Ask Ron (E-24)	SRA
Moore, Susan	Mr. Crown Meets Mr. Who (E-25)	SRA
Moore, Susan	Cowboy Buck (E-6)	SRA
Moore, Susan	The Boy and the Kitten (E-27)	SRA

AUTHOR	TITLE	PUBLISHER
Moore, Susan	Where Is It? (E-28)	SRA
Moore, Susan	The Little Hawk (E-29)	SRA
Moore, Susan	Peppy and the Straw Rooster (E-30)	SRA
Moore, Susan	Jack's Electric Train (E-31)	SRA
Moore, Susan	The Caterpillar (E-32)	SRA
Moore, Susan	The Cameraman (E-33)	SRA
Moore, Susan	A Job for the Camel (E-34)	SRA
Moore, Susan	The Chickadee's Ugly Pal (E-35)	SRA
Moore, Susan	Treat Time (E-36)	SRA
Moore, Susan	A Fish Tale (E-37)	SRA
Moore, Susan	Children Are No Bother? (E-38)	SRA
Moore, Susan	The Eagle Who Wouldn't Fly (E-39)	SRA
Moore, Susan	The Green Pop Bottle (E-40)	SRA
Moore, Susan	Kate's Flower Shop (E-41)	SRA
Moore, Susan	A Moment of Fun (E-42)	SRA
Moore, Susan	April, the Ladylike Alligater (E-43)	SRA
Moore, Susan	Tiger Lily and the Pony (E-44)	SRA
Moore, Susan	Spring Comes to Yang (E-45)	SRA
Moore, Susan	Take Care (F-1)	SRA
Moore, Susan	The Whale Hunt (F-2)	SRA
Moore, Susan	Jenny's Party (F-3)	SRA
Moore, Susan	The Man with Red Hair (F-4)	SRA
Moore, Susan	The Sparrow and the Airplane (F-5)	SRA
Moore, Susan	Henry's Dwarf (F-6)	SRA
Moore, Susan	A Kite Sees the World (F-7)	SRA
Moore, Susan	Snowstorm on the Trail (F-8)	SRA
Moore, Susan	Winter (F-9)	SRA
Moore, Susan	Rosemary and the Purple Purse (F-10)	SRA
Moore, Susan	The Best Tuesday (F-11)	SRA
Moore, Susan	The End of the Red Thread (F-12)	SRA
Moore, Susan	Lost and Found (F-13)	SRA
Moore, Susan	Chuck's Floppy Shoes (F-14)	SRA
Moore, Susan	Tall Boy (F-15)	SRA
Moore, Susan	Salty, the Sad Sardine (F-16)	SRA
Moore, Susan	The Blue Thief (F-17)	SRA

AUTHOR	TITLE	PUBLISHER
Moore, Susan	The Goose Who Loved to Read (F-18)	SRA
Moore, Susan	Sooner or Later (F-19)	SRA
Moore, Susan	Bare Foot, the Indian Boy (F-20)	SRA
Moore, Susan	Maxwell Maddox (F-21)	SRA
Moore, Susan	The Train Trip (F-22)	SRA
Moore, Susan	A Husband for the Princess (F-23)	SRA
Moore, Susan	Bruce and the Red Bicycle (F-24)	SRA
Moore, Susan	The Fireplace (F-25)	SRA
Moore, Susan	Parking a Horse (F-26)	SRA
Moore, Susan	Toughy Announces the News (F-27)	SRA
Moore, Susan	A Lesson in Checkers (F-28)	SRA
Moore, Susan	The Magic Package (F-29)	SRA
Moore, Susan	General Gee Steals the Show (F-30)	SRA
Moore, Susan	The Lion in the Garden (F-31)	SRA
Moore, Susan	It Can Be Anything! (F-32)	SRA
Moore, Susan	The Boy Who Found Antumn (F-33)	SRA
Moore, Susan	Bring Back Snappy (F-34)	SRA
Moore, Susan	The Wise Man and the Well (F-35)	SRA
Moore, Susan	A Home for Mrs. Wren (F-36)	SRA
Moore, Susan	Dave Goes to A Barn Raising (F-37)	SRA
Moore, Susan	One Leg of a Horse (F-38)	SRA
Moore, Susan	My Collection (F-39)	SRA
Moore, Susan	Benny Helps a Friend (F-40)	SRA
Murphy	The Animal Hat Shop	Follett
Murphy	The Roly Poly Cookie	Follett
Rasmussen	A Pig Can Jig	SRA
Rasmussen	A Hen in a Fox's Den	SRA
Rasmussen	Six Ducks on a Pond	SRA
Rasmussen	A King on a Swing	SRA
Rasmussen	Kittens and Children	SRA
Rasmussen	The Purple Turtle	SRA
Seuss	Hop on Pop	Random House
Seuss	Green Eggs and Ham	Random House
Seuss	The Cat in the Hat	Random House
Seuss	The Cat in the Hat Comes Back	Random House

AUTHOR	TITLE	PUBLISHER
Stratemeyer	Tuggy	Harper-Row
Stratemeyer	Frog Fun	Harper-Row
Stratemeyer	Pepper	Harper-Row
Stratemeyer	Six in a Mix	Harper-Row
Stratemeyer	It Happens on a Ranch	Harper-Row
Sullivan	Pins and Pans (1)	Webster/McGraw Hill
Sullivan	The Bag in the Sand (2)	Webster/McGraw Hill
Sullivan	The Red Mitten (3)	Webster/McGraw Hill
Sullivan	The Witch and the Bat (4)	Webster/McGraw Hill
Sullivan	Pick a Pet (5)	Webster/McGraw Hill
Sullivan	Jars and Jars of Jam (6)	Webster/McGraw Hill
Sullivan	The Starship (7)	Webster/McGraw Hill
Sullivan	On the Farm (8)	Webster/McGraw Hill
Sullivan	Patty's Birthday Party (9)	Webster/McGraw Hill
Sullivan	The Baker's Wife (10)	Webster/McGraw Hill
Sullivan	The Big Dragon Hunt (11)	Webster/McGraw Hill
Sullivan	A Visit to Toyland (12)	Webster/McGraw Hill
Sullivan	Come to the Party (13)	Webster/McGraw Hill
Sullivan	Dick and Richard (14)	Webster/McGraw Hill
Sullivan	Dan's Hat (1B)	Webster/McGraw Hill
Sullivan	Ben and Pal (2B)	Webster/McGraw Hill
Sullivan	The Land in the West (3B)	Webster/McGraw Hill
Sullivan	Ned's Pets (4B)	Webster/McGraw Hill
Sullivan	The King's Helper (5B)	Webster/McGraw Hill
Sullivan	The Red Ball (6B)	Webster/McGraw Hill
Sullivan	In the Garden (7B)	Webster/McGraw Hill
Sullivan	Yes I Can (1A)	Webster/McGraw Hill
Sullivan	The Sandman (2A)	Webster/McGraw Hill
Sullivan	The Milkman (3A)	Webster/McGraw Hill
Sullivan	That Kitten! (4A)	Webster/McGraw Hill
Sullivan	Sam's Band (5A)	Webster/McGraw Hill
Sullivan	Things I Think Of (6A)	Webster/McGraw Hill
Sullivan	Presents and Picnics (7A)	Webster/McGraw Hill
Sullivan	At the Market (8A)	Webster/McGraw Hill
Sullivan	Kitty and James (9A)	Webster/McGraw Hill

AUTHOR	TITLE	PUBLISHER
Sullivan	The Big Race (10A)	Webster/McGraw Hill
Sullivan	Animals Here and There (11A)	Webster/McGraw Hill
Sullivan	Far North to the Big City (12A)	Webster/McGraw Hill
Sullivan	Dan and Greta (13A)	Webster/McGraw Hill
Sullivan	Can You Tell a Story (14A)	Webster/McGraw Hill
Sullivan	The Spy in the Sky (1)	BRL
Sullivan	The Fox and the Cat (1A)	BRL
Sullivan	The Map in the Pan (1B)	BRL
Sullivan	The Camp by the Pond (2)	BRL
Sullivan	The X Y X Gang (2A)	BRL
Sullivan	Pam and the Hat (2B)	BRL
Sullivan	Bad Luck Sam (3)	BRL
Sullivan	The Big Lab (3A)	BRL
Sullivan	The Best Club (3B)	BRL
Sullivan	The End of the Bad Luck (4)	BRL
Sullivan	The Spy Hunt (4A)	BRL
Sullivan	A Job on a Jet (4B)	BRL
Sullivan	A Day on the Farm (5)	BRL
Sullivan	The Old TV Set (5A)	BRL
Sullivan	The Big Show (5B)	BRL
Sullivan	Tom Finds a Job (6)	BRL
Sullivan	The First Man on the Moon (6A)	BRL
Sullivan	The Cave Full of Gold (6B)	BRL
Sullivan	The First Wheel (7)	BRL
Sullivan	The TV World (7A)	BRL
Sullivan	The Chase (7B)	BRL
Sullivan	Fred and His Fears (8)	BRL
Sullivan	A Girl, a Book, and a Car (8A)	BRL
Sullivan	To Brag Is to Boast (8B)	BRL

APPENDIX D

Samples of Directions Given to Scorers of Printed Tests

Gates-MacGinitie Reading Test
Gilmore Oral Reading Test
San Diego Inventory of Reading Attitudes

DIRECTIONS FOR SCORING THE GATES-MacGINITIE PRIMARY A

Please read carefully and follow rigidly.

The raw score on either the Vocabulary or Comprehension test is the total number of items for which the child chooses the correct answer. The form or neatness of the mark is not considered. However, *more than one mark* for a numbered item makes the entire item wrong *unless* all but one marked item has been scribbled out. (The children were told they could change their minds by scribbling over their mistakes and then marking another selection.)

Scoring keys are included and contain directions for their use.

Enter the raw score (number right for a section, Vocabulary or Comprehension) on the front of the test booklet. Compute the Standard Score, Percentile, and Grade Score by referring to page 7, Tables I-IV, of the Teacher's Manual. For Tables I and III, look in the column for Grade Level 1.5 (Feb.).

With black or blue ballpoint pen, carefully transfer each child's raw score, standard score, and grade score for Vocabulary and Comprehension to the enclosed record sheet under the columns labeled Gates-MacGinitie Primary AI at the top.

FIRST GRADE PROJECT
GILMORE ORAL READING TEST – SCORING DIRECTIONS
(See page 14 of Manual of Directions for Sample)

The Accuracy, Comprehension, and Rate data should be transferred to the "Test Summary" on the title page of the Record Blank according to the directions below.

Computing the Accuracy Score

1. In the *Errors* column on the title page record the number of errors made on each paragraph read.
2. In the *10 Minus No. Errors* column, record 10 minus the number of errors made in each paragraph. (The highest possible score for any paragraph is then 10 if it is read without error. 4 errors would be 10-4, or 6.) There can be no negative scores listed. 12 errors would be 10-12, for which you would record 0.
3. The sum of the scores in the *10 Minus No. Error* column gives the Accurary Score entered in the large square at the bottom.

Computing the Comprehension Score

1. A comprehension score for *each* paragraph is recorded in the *Number Right* column.
2. One point is given for each question answered correctly. (The highest possible Comprehension score for any paragraph is 5.)

3.* In the paragraph *immediately above* the pupil's ceiling level, credit is given for answering one less question than he received at his ceiling level. On the next higher paragraph, he is given credit for answering one less question than on the preceding paragraph and so on.

4. The pupil's comprehension score is the sum of the entries in the *Number Right* column. It is entered in the large square at the bottom.

Computing the Rate Score

1. In the *Time in Seconds* column, record the time required for reading each paragraph.

2.* Add the number of words in all paragraphs read *EXCEPT* the *ceiling* paragraph and record in square 1 below.

3.* Add the times for all paragraphs read *EXCEPT* the *ceiling* paragraph and record in square 2 below.

4. To compute the rate in seconds, divide the number of words read (1) by the time in seconds (2). (Carry out to two decimal places and round off to one. Multiply the answer by 60 to get Words per Minute.)

Obtaining Grade Equivalents Corresponding to Raw Scores

1. For *Accuracy* see Table 8 on page 19 of Manual of Directions. FORM A
2. For *Comprehension* see Table 9 on page 20 of Manual of Directions. FORM A

Totaling Errors

Simply total all errors of a specified type made on *all* paragraphs read, and record on the Research Project Record Sheet.

*Important. Read Carefully

SAN DIEGO INVENTORY OF READING ATTITUDES
SCORING DIRECTIONS

KEY: Questions 2, 9, 19, 20 should be answered *NO:* the remaining questions should be answered *YES.*

Mark the correct responses according to the key above. Record the raw score (total correct responses) and the stanine (from table below) on the front of each child's Inventory.

Raw Scores	Stanines
25	9
23-24	8
21-22	7
19-20	6
16-18	5
13-15	4
10-12	3
7-9	2
0-6	1

Record the child's name, Raw Score, and Stanine on the *classroom record sheet.*

APPENDIX E

Motivation Sheet for Creative Writing Sample I

MOTIVATION SHEET FOR CREATIVE WRITING SAMPLE I

(Say words written in capitals)

Pass out writing paper to the children and check to see that they all have sharp pencils. Have each write his name on the top line.

DID YOU KNOW THERE ARE STORES YOU CAN GO TO, TO BUY JUST ANY KIND OF ANIMAL YOU'VE EVER WANTED FOR A PET? WELL, THERE ARE, AND THEY'RE CALLED PET STORES OR ANIMAL STORES. I'M GOING TO READ YOU A POEM ABOUT AN ANIMAL STORE, AND WHILE I'M READING I WANT YOU TO LISTEN AND THEN THINK ABOUT THE ANIMAL *YOU* WOULD BUY IF YOU WENT TO AN ANIMAL STORE FOR JUST ANY PET YOU WANTED.

THE ANIMAL STORE (by Rachel Field)

IF I HAD A HUNDRED DOLLARS TO SPEND,
OR MAYBE A LITTLE MORE,
I'D HURRY AS FAST AS MY LEGS WOULD GO
STRAIGHT TO THE ANIMAL STORE.

I WOULDN'T SAY, "HOW MUCH FOR THIS OR THAT?"
"WHAT KIND OF A DOG IS HE?"
I'D BUY AS MANY AS ROLLED AN EYE,
OR WAGGED A TAIL AT ME!

I'D TAKE THE HOUND WITH THE DROOPING EARS
THAT SITS BY HIMSELF ALONE;
COCKERS AND CAIRNS AND WOBBLY PUPS
FOR TO BE MY VERY OWN.

I MIGHT BUY A PARROT ALL RED AND GREEN,
AND THE MONKEY I SAW BEFORE,
IF I HAD A HUNDRED DOLLARS TO SPEND,
OR MAYBE A LITTLE MORE.

NOW, CLOSE YOUR EYES FOR JUST A MINUTE AND THINK ABOUT WHAT ANIMAL YOU WOULD BUY IF YOU WENT TO THE

ANIMAL STORE. IT COULD BE ANY PET–BIG OR LITTLE, WILD OR TAME, OR EVEN STRANGE AND UNUSUAL. (Allow about one minute for them to think.) NOW, I WANT YOU TO WRITE ALL ABOUT THAT ANIMAL ON THE PAPER YOU HAVE ON YOUR DESK. I WON'T TELL YOU HOW TO SPELL ANY WORDS. YOU SPELL THEM THE BEST YOU CAN YOURSELF. I WANT TO KNOW ABOUT YOUR ANIMAL, SO WRITE ALL YOU CAN THINK OF TO TELL ME IN YOUR STORY. YOU MAY BEGIN NOW, AND REMEMBER, SPELL THE WORDS AS BEST YOU CAN. I WON'T TELL YOU HOW. WHEN YOU ARE FINISHED, BRING YOUR PAPERS UP TO ME.

Allow 20 minutes of writing time. In order that the stories will be clear to the scorer, as each child finishes ask him to read his paper to you. In red pencil, write in unclear words above the words written. The teacher should be asked to assist you in hearing the stories, so instruct her in writing unclear words. Call in all papers, finished or not, at the end of 20 minutes.

APPENDIX F

Scoring Directions for Written Composition

SCORING DIRECTIONS FOR WRITTEN COMPOSITION

Measures To Be Obtained

1. No. of words attempted.

2. $$\frac{\text{No. of phonetically acceptable words}}{\text{No. of words attempted}} \times 100 = \%\ \text{phonetically acceptable words}$$

3. $$\frac{\text{Total No. independent clauses} \div \text{dependent clauses}}{\text{No. independent clauses}} = \text{Dependent Clause Index}$$

4. $$\frac{\text{No. of words}}{\text{No. of T-Units}} = \text{Mean T-Unit Length}$$

Directions and Criteria for Correction

1. Count and record in proper column of record sheet the number of words attempted. The words written in red were added by the examiner to clarify for you the intent of the child. When they appear in parentheses, they are words added by the child in the reading of the story and *should not* be included in a count of attempted words.

2. Compute the formula above (2) and record in proper column of record sheet. Words will be accepted as "phonetically acceptable" if:
    a. They are correct according to conventional English spelling.
    b. They are homonyms of the word indicated by the context and as such are correct according to conventional English spelling.
    E.g., *to* for to, too, two
    wood for wood or would
    c. They contain a close approximation to *each* phoneme of the word indicated by the context. (Keeping in mind possible dialectal or grammatical variations and giving the benefit of the doubt when there is serious question.)

E.g., wen for when	acceptable
anamal or animal	acceptable
runned or runed for ran	acceptable
wat for want	not acceptable

3. Compute the formula above (3) and record in proper column of record sheet.
    a. An independent clause is defined as a clause which can stand by itself to make a sentence. For the purposes of this study they do not have to be grammatically "correct." They are acceptable with missing function words or even missing verbs if the intent seems to be formation of an independent clause.
    E.g., Look the monkey. (One clause)
    It not bite you. (One clause)
    Look the monkey and it not bite you. (Two clauses)
    b. A dependent clause either (1) helps to compose expansions of smaller grammatical units, or (2) is equivalent to single nominals, adjectivals, or adverbials, or (3) both. (Sledd) For the purposes of this study, they do not have to be grammatically "correct." They are acceptable with missing words if intent seems to be formation of a dependent clause.
    E.g., *After I buy a dog,* I will buy a cat.
    The cat I *buyed* is black.

4. Compute the formula above (4) and record in proper column of record sheet. T-Units (Minimal Terminal Units–Hunt) are the shortest units into which a piece of discourse can be cut without leaving sentence fragments as residue. Each is one independent clause ÷ whatever dependent clause(s) are attached. They intervene between the independent clause and the compound sentence.
E.g., I want a pet can. One T-Unit
After I buy a cat, I want a dog. One T-Unit
I will buy a cat and I will buy a dog. Two T-Units

SCORING EXAMPLE

one one Peek
My pets are kittens. The bigast *oan* is Bo. The smallist *oan* is *penk.*
They drink
Thay *dink* milk. Thay plat a lot.

Italicized words were judged not phonetically acceptable.

1. 21 words attempted.
2. $\frac{17}{21} \times 100 = 809 = 81\%$ (rounded)

3. $\frac{5 \div 0}{5} = 1$

4. $\frac{21}{5} = 4.2$ (to one decimal)

RECORDING EXAMPLE

	Written Comp.				
	No. Words Attemp.	Phonet. %	Clause Index	T-Unit Length	
17. Jay Smith	21	.81	1	4.2	

APPENDIX G

Directions and Scoring Sheet for Free-Time Motivational Study Developed by LaVonne Norin

It is being requested that a sample of the teachers in the first-grade reading study conduct and record a 10-minute observation twice a week for a period of three weeks in the month of February as follows:

1. Give the students a 10-minute activity period twice a week. Instruct the students in the following manner: "During the next 10 minutes you may read, write or draw. You may read a book of your choice, or write a story or letter, or draw pictures and color them. You may choose only one of these activities for the 10-minute period. Once you have begun an activity you may not change it during the period. However, the next time we have an activity period you may choose a different activity if you wish."

2. List all the students in your class on the form provided. For those students who select reading as an activity, mark a check (✓) in the space provided by the student's name. If a student is absent during this time, please mark an A in the space provided.

NOTE: The selection of the time for the 10-minute activity should be kept constant within the class. It is suggested that one period of time each week is given in the morning and the second period of time given in the afternoon, each on different days.

SCHOOL ____________________ TEACHER ____________________

	Morning session	After-noon session	Morning session	After-noon session	Morning session	After-noon session
STUDENT NAME	DATE	DATE	DATE	DATE	DATE	DATE